PATENT COOPERATION TREAT

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Commissioner
US Department of Commerce
United States Patent and Trademark
Office, PCT
2011 South Clark Place Room
CP2/5C24
Arlington, VA 22202
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 30 October 2000 (30.10.00)	
International application No. PCT/US00/04925	**Applicant's or agent's file reference** 00786/371WO2
International filing date (day/month/year) 24 February 2000 (24.02.00)	**Priority date** (day/month/year) 24 February 1999 (24.02.99)
Applicant SEED, Brian et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

05 September 2000 (05.09.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer Manu Berrod
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 00786-371WO2	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US00/04925	International filing date *(day/month/year)* 24 FEBRUARY 2000	Priority date *(day/month/year)* 24 FEBRUARY 1999
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant THE GENERAL HOSPITAL CORPORATION		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets.

   ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of report with regard to novelty, inventive step or industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
05 SEPTEMBER 2000	13 JUNE 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI Telephone No. (703) [illegible]

TERRY J. DEY
PATENT LEGAL SPECIALIST
TECHNOLOGY CENTER 1600

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/04925

I. Basis of the report

1. With regard to the **elements** of the international application:*

[X] the international application as originally filed

[X] the description:
pages 1-29 , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the claims:
pages 30-37 , as originally filed
pages NONE , as amended (together with any statement) under Article 19
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the drawings:
pages 1-8 , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

[X] the sequence listing part of the description:
pages NONE , as originally filed
pages NONE , filed with the demand
pages NONE , filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language _____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

[ ] contained in the international application in printed form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages NONE

[X] the claims, Nos. NONE

[X] the drawings, sheets/~~fig~~ NONE

5. [ ] This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

** Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

***Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/04925

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been and will not be examined in respect of:

☐ the entire international application.

☒ claims Nos. 5-6, 17-18, 24-43

because:

☐ the said international application, or the said claim Nos. _ relate to the following subject matter which does not require international preliminary examination *(specify)*.

☐ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. _ are so unclear that no meaningful opinion could be formed *(specify)*.

☐ the claims, or said claims Nos. _ are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for said claims Nos. 5-6, 17-18, 24-43 .

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/04925

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1-4, 7-16, 19-23	YES
	Claims	NONE	NO
Inventive Step (IS)	Claims	NONE	YES
	Claims	1-4, 7-16, 19-23	NO
Industrial Applicability (IA)	Claims	1-4, 7-16, 19-23	YES
	Claims	NONE	NO

2. citations and explanations (Rule 70.7)

Claims 1-4, 7-16, 19-23 lack an inventive step under PCT Article 33(3) as being obvious over either Birnbaum (Journal of Virology, April 1994, Vol 68, No. 4, pages 2521-2528) or Clem et al (Science, vol. 254, November 1991, pages 13881390)in view of Zolotukin et al (US Patent 5,874,304).

Birnbaum et al teach an apoptosis -inhibiting gene from s nuclear polyhedrolsis virus encoding a polypeptide with Cys/His sequence motifs. The reference teaches that two different baculovirus genes are known to be able to block apoptosis triggered upon infection of Spodoptera cells with p35 mutants of the insect baculovirus. the reference teaches that using a genetic complementation assay to identify additional genes which inhibit apoptosis during infection with p35 mutant, Birnbaum et al have isolated a gene OpNPV that was able to functionally substitute for AcMNPV p35. The reference teaches a genetic complementation assay to identify apoptosis blockin genes in which SF-21 cells were cotransfected with vAcAnh and the test baculovirus DNA and transfected cells monitored.

Clem et al teach prevention of Apoptosis by a Baculovirus Gene during Infection of Insect cells. Clam et al have identif ied annihilator (vAcAnh) viral mutant was identified from expression vectors. The mutant caused premature death of Spodoptera frugiperda cells.

Either Birnbaum et al or Clem et al have taught the method of identification of the gene or the polypeptide involved in apoptosis using the a reporter gene expression. However, Zolotukhin et al teach humanized green fluorescent protein genes and methods of use of the protein in several assays. The reference teaches that the expression vectors may comprise a multiple cloning sites that is operatively positioned downstream from gfp gene sequence, and these vectors are useful in addition to the uses in creating C terminal fusion proteins by cloning a second protein encoding DNA segment into the multiple cloning site so that it is contiguous and in frame with the gfp sequence. The reference teaches that the recombinant host cells will express GFP to produce encoded GFP protein, preferably in amount sufficient to allow GFP detection by its fluorescence. The reference teaches that the expression vectors comprising a GFP (Continued on Supplemental Sheet.)

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

CLASSIFICATION:
The International Patent Classification (IPC) and/or the National classification are as listed below:
IPC(7): C12Q 1/00, 1/70, 1/66; C12N 7/00, 5/00, 5/02; G01N 33/554; A61K 38/00; C07H 21/02, 21/04, 21/00 and US Cl.: 435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):
gene operatively linked to a selected gene, and the fusion protein being produced in amount sufficient to allow cell detection by detecting the green fluorescence of GFP. Thus, it would have been obvious to use the GFP protein as a reporter gene in the method of identifying the gene involved in the apoptosis.

Claims 1-4, 7-16, 19-23 meet the criteria set out in PCT Article 33(2) and (4), because the prior art does not teach the method of identifying the a polypeptide which increases or decreases gene expression from a promoter using a GFP reporter gene operably linked to said promoter.

---------- NEW CITATIONS ----------
NONE

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4, 7-16, 19-23

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Ref. 09/919/635

MOLECULAR AND CELLULAR BIOLOGY, Feb. 1995, p. 704–710
0270-7306/95/$04.00+0

Vol. 15, No. 2

Expression Cloning of Oncogenes by Retroviral Transfer of cDNA Libraries

IAN WHITEHEAD, HEATHER KIRK, AND ROBERT KAY*

Department of Medical Genetics, University of British Columbia, and Terry Fox Laboratory, British Columbia Cancer Agency, Vancouver, British Columbia, Canada V5Z 4E6

Received 23 June 1994/Returned for modification 1 September 1994/Accepted 27 October 1994

A cDNA library transfer system based on retroviral vectors has been developed for expression cloning in mammalian cells. The use of retroviral vectors results in stable cDNA transfer efficiencies which are at least 100-fold higher than those achieved by transfection and therefore enables the transfer and functional screening of very large libraries. In our initial application of retroviral transfer of cDNA libraries, we have selected for cDNAs which induce oncogenic transformation of NIH 3T3 fibroblasts, as measured by loss of contact inhibition of proliferation. Nineteen different transforming cDNAs were isolated from a total of 300,000 transferred cDNA clones. Three of these cDNAs were derived from known oncogenes (*raf-1*, *lck*, and *ect2*), while nine others were derived from genes which had been cloned previously but were not known to have the ability to transform fibroblasts (β-catenin, thrombin receptor, phospholipase C-γ_2 and Spi-2 protease inhibitor genes). The Spi-2 cDNA was expressed in antisense orientation and therefore is likely to act as an inhibitor of an endogenous transformation suppressor. Seven novel cDNAs with transforming activities, including those for three new members of the CDC24 family of guanine nucleotide exchange factors, were also cloned from the retroviral cDNA libraries. Retroviral transfer of libraries should be generally useful for cloning cDNAs which confer selectable phenotypes on many different types of mammalian cells.

Cellular oncogenes or proto-oncogenes can be cloned by selecting for their ability to confer the phenotype of deregulated growth on cells in which they are expressed. This was first done by transfecting fragmented genomic DNA into cell lines such as NIH 3T3 fibroblasts which are susceptible to single-hit oncogenic transformation (13, 15, 31, 36, 40, 46). Subsequently, stable transfection of these cells with cDNA libraries in plasmid or phage expression vectors has been used as an alternative approach, to avoid the severe difficulties which have been encountered in recovering and analyzing oncogenes present in transfected genomic DNA (8, 9, 32–34). Despite the theoretical advantages of working with transfected cDNA expression libraries, only a small fraction of the many currently known oncogenes have been cloned in this way. The most significant limitation in the use of cDNA library transfer for cloning oncogenes has been the low efficiencies of cDNA transfer and expression which can be achieved by stable transfection methods. It is difficult to generate more than a few tens of thousands of transfectants in which cDNA clones are being expressed at adequate levels, but a comprehensive screening of a mammalian cDNA library demands the transfer and expression of several million clones. Therefore, the small number of oncogenes which have been cloned so far from transferred libraries is unlikely to be due to exhaustion of the pool of cDNAs with oncogenic potential but rather is a result of a failure to transfer, and thus detect, most such cDNAs.

In contrast to deliberately constructed cDNA libraries, natural populations of retroviruses have served as a prolific source of oncogenes, through their ability to incorporate and mutationally activate cDNA copies of host cell mRNAs and then to transfer them to recipient cells by infection (5). Artificial retroviral vectors with equivalent abilities to transfer and express cDNAs in mammalian cells (35) could be ideally suited to screening libraries for cDNAs conferring selectable phenotypes such as deregulated cell growth. Previous reports have indicated the potential for using retroviruses to clone cDNAs conferring selectable phenotypes, as demonstrated by the isolation of cDNAs encoding thymidine kinase (TK) via rescue of TK^- fibroblasts (39) and cDNAs encoding interleukin-3 and granulocyte-macrophage colony-stimulating factor via conversion of hemopoietic cell lines to cytokine-independent growth (42). This report describes the development of an efficient system for cDNA library transfer and expression by retroviral vectors and its application to the cloning of a large number of cDNAs which induce oncogenic transformation of murine fibroblasts.

MATERIALS AND METHODS

pCTV vectors. pCTV1 was constructed from the following elements: (i) an *Nhe*I-*Kpn*I fragment containing the U3 region and part of the R region of Moloney murine leukemia virus (MMLV), excised from the JZen retroviral vector (22); (ii) a *Kpn*I-*Xho*I fragment extending from the R region to base 1035 in the MMLV genomic sequence (GenBank accession number J02255), derived by PCR from the pBabeHygro retroviral vector (38); (iii) an *Xho*I-*Cla*I fragment containing a synthetic cloning site and an adjacent *supF* gene (43), derived by PCR from the pAX114 expression vector (24); (iv) a *Cla*I-*Bam*HI fragment from a subclone containing the myeloproliferative sarcoma virus (MPSV)-MMLV hybrid 3′ long terminal repeat (LTR) from JZen along with 110 bp of uncharacterized genomic DNA sequence downstream of the LTR; and (v) a *Bgl*II-*Spe*I fragment from pAX114, containing the β-globin and simian virus 40 (SV40) polyadenylation sites, the SV40 origin of replication, and the origin of replication from pUC8.

pCTV3 was constructed by inserting a *Cla*I-*Acc*I fragment containing a hygromycin resistance gene into the *Cla*I site of pCTV1. This gene consists of a synthetic polyomavirus enhancer and a herpes simplex virus TK promoter (49) attached to a hygromycin phosphohydrolase gene (3) with a modified translation initiation sequence.

pCTV1B and pCTV3B were derived by inserting the 380-bp *Bst*YI fragment containing the *Bst*XI stuffer fragment of pCDM8 (44) between the *Sal*I sites of pCTV1 or pCTV3.

pCTV3K was derived from pCTV3 by replacement of the *supF* gene with the P1 promoter of pBR322 and the neomycin phosphotransferase coding region of Tn*5* (1).

* Corresponding author. Mailing address: Terry Fox Laboratory, BC Cancer Agency, 601 W. 10th Ave., Vancouver, BC, Canada V5Z 4E6. Phone: (604) 877-6070. Fax: (604) 877-0712. Electronic mail address: robert@terryfox.ubc.ca.

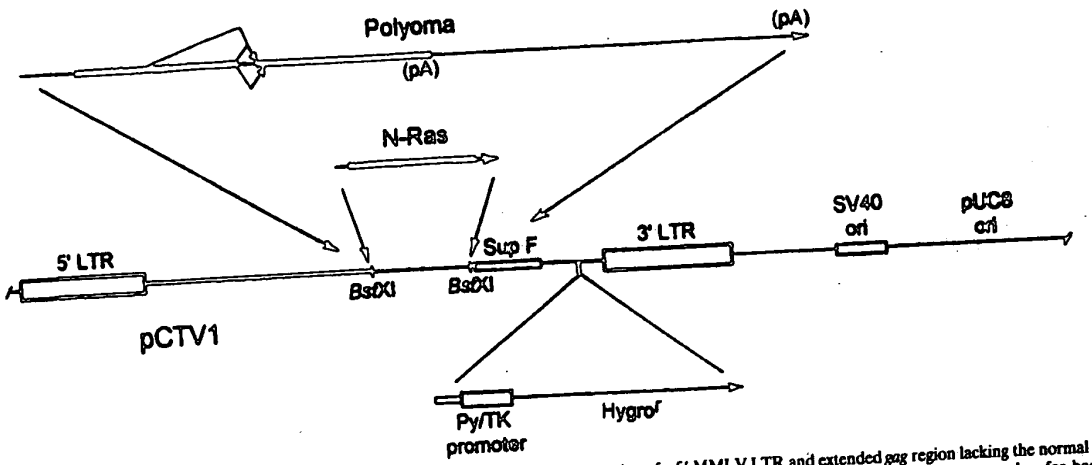


FIG. 1. Structures of retroviral vectors for cDNA library transfer. The pCTV1 vector consists of a 5′ MMLV LTR and extended *gag* region lacking the normal splice donor and initiation codon, a polylinker with a *Bst*XI stuffer which is displaced by cDNA insertion, a *supF* gene which provides a compact marker for bacterial transformation as well as a marker for cDNA recovery, a composite MPSV-MMLV 3′ LTR, and replication origins from pUC8 and SV40. The total size of pCTV1 is 4,350 bp. The pCTV3 vector is identical except for the presence of a 1,210-bp hygromycin phosphohydrolase gene driven by a polyomavirus (Py) enhancer and herpes simplex virus TK promoter. An exact description of the structure and construction of these vectors is available on request. After viral transmission, the region including and between the two LTRs becomes integrated into the recipient cell's genome as a provirus, with both 5′ and 3′ LTRs now being derived from the composite MPSV-MMLV LTR. Transcription of the provirus is initiated within the 5′ LTR, and polyadenylation occurs within the 3′ LTR. The N-Ras cDNA and polyomavirus early genes used to test for cDNA transfer and expression efficiencies are shown above the vector, with the alternative splicing and potential polyadenylation signals (pA) marked on the polyomavirus construct.

cDNA constructs. The CD24 cDNA used for quantifying cDNA transfer and expression was the 950-bp clone described previously (25). The N-Ras cDNA was amplified by PCR from a cDNA library prepared from the SK-N-SH neuroblastoma cell line (13). This cDNA has a Gln-61-to-Lys mutation which results in oncogenic activation. The polyomavirus early region construct extended from 153 to 2926 in the polyomavirus genomic sequence (GenBank accession number J02288).

Construction of cDNA libraries in pCTV vectors. Poly(A) RNA was prepared from the murine hemopoietic cell lines GM979 (7), 32D (17), and B6SUtA$_1$ (17) by guanidinium isothiocyanate lysis and oligo(dT) chromatography as described previously (12, 21). cDNA was synthesized from the poly(A) RNA by using MMLV reverse transcriptase (Life Technologies, Gaithersburg, Md.) and random sequence primers, with the procedures and reagents recommended by the supplier. 5′ phosphorylated *Bst*XI adapters (TCAGTTACTCAGG and CCTGA GTAACTGACACA) were ligated to the double-stranded cDNA, which was then size fractionated by agarose gel electrophoresis. Twenty nanograms of *Bst*XI-digested and dephosphorylated pCTV3B or pCTV1B plasmid DNA was ligated with an equimolar amount of the size-fractionated cDNA in 10 μl of 25 mM Tris-Cl (pH 7.8)–5 mM $MgCl_2$–1 mM dithiothreitol–1 mM ATP–0.5 U of T4 DNA ligase (Life Technologies); the mixture was incubated at 16°C for 3 h and then for 20 min at 72°C. The ligation reaction was used to transform *Escherichia coli* MC1061/p3 (45) by electroporation, using the procedures described in the operating manual for the GenePulser electroporation apparatus (BioRad, Richmond, Calif.). After a 90-min incubation at 37°C, the transformed bacteria were plated in soft agar, incubated at 37°C for 20 h, and then recovered from the agar as described previously (50). Plasmid DNA was prepared from the pooled bacterial colonies by the alkaline lysis procedure (4), digested with RNase A and RNase T_1, and precipitated with ammonium acetate and ethanol.

Transfection of packaging cells and infection of NIH 3T3 cells. Six-centimeter-diameter culture dishes containing 1.5×10^6 GP+E-86 cells (29) were washed twice with Dulbecco's modified Eagle's medium (DME; Stem Cell Technologies, Vancouver, British Columbia, Canada) containing 20 mM *N*-2-hydroxyethylpiperazine-*N*-2-ethanesulfonic acid (HEPES; pH 7.2); 2 ml of DME-HEPES containing 2 μg of plasmid DNA and 200 μg of DEAE-dextran (molecular weight, 500,000; Pharmacia) per ml was then added, and the cells were incubated for 60 min at 37°C. The cells were washed twice with DME-HEPES containing 10% calf serum and incubated for 3 h in DME-HEPES containing 10% calf serum and 200 μM chloroquine, after which the medium was replaced with DME–10% calf serum; 45 h later, the culture medium was replaced with 4 ml of DME–10% calf serum containing 5 mM sodium butyrate to enhance transcription from the transfected plasmids (16). The medium containing secreted viruses was collected 18 h later and filtered, and Polybrene was added to a concentration of 10 μg/ml. This viral supernatant was then added to a 10-cm-diameter dish containing 5×10^5 NIH 3T3 cells. After a 16-h incubation at 37°C, the viral supernatant was replaced with DME–10% calf serum; the NIH 3T3 culture was maintained with twice-weekly feedings of this medium for 3 to 4 weeks to allow the emergence of transformed foci.

PCR recovery of proviral cDNA inserts. Transformed cells were plucked from infected NIH 3T3 cell cultures and expanded to 10^6 cells, and genomic DNA was prepared by proteinase K digestion, phenol extraction, and ethanol precipitation. Fifty-microliter PCR reaction mixes contained 300 μg of genomic DNA, 100 ng of primers annealing to retroviral vector sequences (5′ primer, CTCACTCCT TCTCTAGCTC; 3′ primer, CACACTGCTCGAATCAAGC), 200 μM each deoxynucleoside triphosphate, 1.25 U of *Pfu* DNA polymerase (Stratagene), and the buffer provided with the polymerase. Thermal cycles were as follows: 1 at 95°C for 60 s; 5 at 95°C for 60 s, 50°C for 30 s, and 72°C for 300 s; and 30 at 95°C for 30 s, 50°C for 30 s, and 72°C for 300 s. After cycling, 150 μl of 1.5 M ammonium acetate was added, and the reaction mix was extracted with chloroform and precipitated with 2 volumes of ethanol. Specific PCR products in a portion of the reaction were identified by Southern blotting, using a probe detecting the coamplified *supF* gene. The amplified DNA was then digested with *Mlu*I and *Bsi*WI, and PCR products were purified by agarose gel electrophoresis, electroelution, and ethanol precipitation. The purified PCR products were ligated with *Mlu*I-*Bsi*WI-digested pCTV3K and used for transformation of *E. coli* MC1061/p3. Plasmids were then isolated from single bacterial clones. Plasmid clones which tested positive for transformation were recloned by bacterial transformation and retested for transformation to ensure that transformation was reproducible and was not due to a contaminating plasmid.

RESULTS AND DISCUSSION

Design of retroviral vectors. A set of retroviral vectors was constructed specifically for cDNA library transfer and expression cloning (Fig. 1). The pCTV vectors were made as compact as possible to maximize stability of cDNA inserts during propagation both as plasmids in *E. coli* and as retroviruses. They contain a pair of *Bst*XI sites separated by a stuffer fragment to allow the use of noncomplementary *Bst*XI adapters for cDNA insertion (45). Inserted cDNAs are included in the genomic viral transcript which initiates in the 5′ LTR. This transcript also contains an extended packaging signal to maximize viral titers (2). The *gag* initiation codon within this region has been removed by mutagenesis to allow efficient translation from initiator codons within inserted cDNAs (38). All splicing signals have been removed from the pCTV vectors to avoid the activation of cryptic splice sites within inserted cDNAs, which

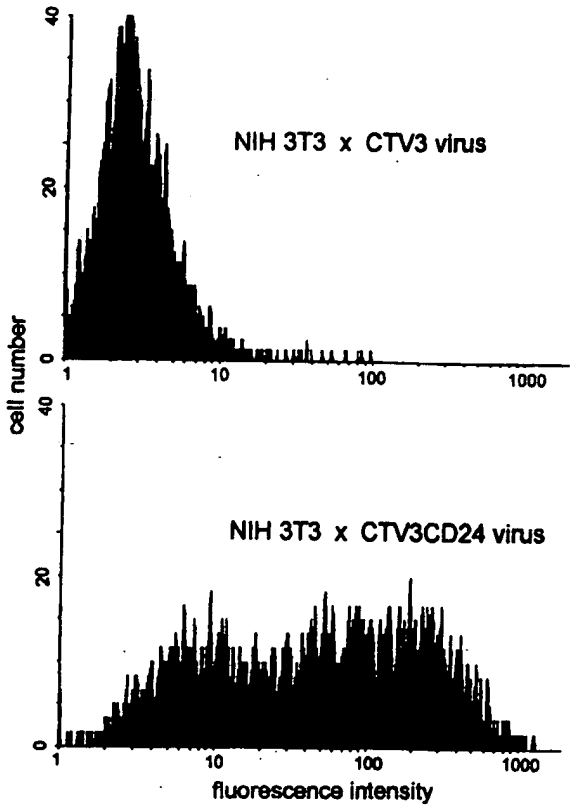


FIG. 2. Determination of cDNA transfer and expression efficiencies. GP+E-86 packaging cells were transfected with pure pCTV3 or pCTV3CD24 plasmid clones, the latter carrying a cDNA encoding the cell surface protein CD24 (25). NIH 3T3 cells were infected with the resulting viral supernatants and 2 days later were stained with an anti-CD24 monoclonal antibody complexed with the fluorochrome phycoerythrin. Expression of CD24, as measured by fluorescence intensity, was analyzed by flow cytometry as described previously (25).

would prevent cDNA transmission. The pCTV vectors also contain SV40 origins of replication so that the libraries can be used for expression cloning strategies dependent on episomal replication in COS cells (24, 45).

Production of retroviruses. Transient transfection of retroviral packaging cell lines was used to achieve the simultaneous, rapid, and proportionate conversion of very large numbers of cDNA clones into cDNA-bearing retroviruses. With a DEAE-dextran-mediated transfection protocol combined with chloroquine and butyrate treatments, titers of about 10^5 infectious viruses per ml, as determined by hygromycin resistance, were obtained with the GP+E-86 packaging line (29). Infection of NIH 3T3 cell cultures with the viral supernatants produced by transient transfection resulted in the transfer of functional viruses to about 70% of the cells, as measured by expression of a cell surface protein encoded by a virus-borne cDNA (Fig. 2). Transient transfection of GP+E-86 cells with pCTV3-based cDNA libraries typically resulted in somewhat lower titers, in the range of 1×10^4 to 4×10^4/ml, possibly reflecting the lower purity of the plasmid DNA obtained directly from bacterial colonies of the plated library. Such titers were still sufficiently high to allow the transfer of 100,000 cDNA clones by using 5 ml of supernatant for the infection of about 300,000 recipient cells. This is in striking contrast to our previous experience with direct transfection of NIH 3T3 and other cell lines, in which case the use of fully optimized transfection procedures (10, 11) and expression vectors (24, 37) resulted in cDNA transfer to only 0.1 to 0.5% of the recipient cells.

TABLE 1. Efficiencies of transfer and expression of transforming cDNAs

Transfected plasmid(s)	Amt of DNA	No. of transformed foci
pCTV1	1 μg	0
pCTV1 + pCTV1/N-Ras	1 μg + 1 ng	~300[a]
pCTV1 + pCTV1/N-Ras	1 μg + 0.1 ng	39
pCTV1 + pCTV1/Polyoma	1 μg + 1 ng	~500[a]
pCTV1 + pCTV1/Polyoma	1 μg + 0.1 ng	62

[a] Estimated focus number, due to high density of foci on dishes.

Detection of rare oncogenic cDNAs within transferred viral populations. One of our initial concerns about using retroviral vectors for cDNA library transfer was the possibility that many cDNA clones, especially longer ones, would be inefficiently transmitted or partially deleted during transmission. This might occur either through interference of the cDNA inserts with the transcription or packaging of the viral genomic RNA or through recombination between different cDNA clones sharing a common packaging cell or virus particle. To competitively test the fidelity of cDNA transmission, we constructed pCTV1 vectors carrying two distinctly different oncogenic cDNAs, i.e., an activated N-Ras cDNA of only 650 bp and a 2,850-bp region encompassing the entire early gene region of polyomavirus. In addition to its much larger size, the polyomavirus clone included alternatively utilized splice sites and two potential polyadenylation signals, features which would be expected to be found in some cDNAs and which have the potential to block formation of the intact genomic RNA needed for transmission. The N-Ras and polyomavirus vectors were highly diluted with pCTV1 vector to mimic a library containing rare oncogenic cDNAs. These DNA mixtures were then used to transfect GP+E-86 cells, and the resulting supernatants containing released virus were used to infect NIH 3T3 cells. The small N-Ras clone and the large and complex polyomavirus clone were transmitted to the NIH 3T3 cells with similar efficiencies, which were roughly equivalent to their abundances within the population of plasmids used for the initial transfection of the packaging cells (Table 1). Large numbers of transformed foci were induced in the infected NIH 3T3 cell cultures even when the abundances of the vectors carrying the N-Ras or polyomavirus clones were only 1:10,000 relative to inert vector. This result suggested that retrovirally transferred cDNA libraries could be used to detect the presence of transforming cDNAs with abundances of less than 1:100,000.

Recovery of retrovirally transferred cDNAs. To facilitate the analysis of the large number of transformed cell clones which were expected to be generated from cDNA library transfers, a PCR strategy was developed for rapidly recovering cDNAs from selected recipient cell clones and immediately reinserting them into the pCTV3 vector (Fig. 3). Several thousand bacterial transformants were obtained from cDNA recovery procedures performed on transformed NIH 3T3 cell clones derived from infection with the 10,000:1 mixture of pCTV1 with pCTV1/N-Ras. The yields of bacterial transformants from recovery procedures performed on the polyomavirus-transformed cell clones were about 20-fold lower, in proportion to the lower yield of PCR amplification of the longer polyomavirus cDNA. All recovered cDNA clones which were examined had retained the original cDNA structure, with the exception in all polyomavirus clones of the loss of the 61-bp intron spe-

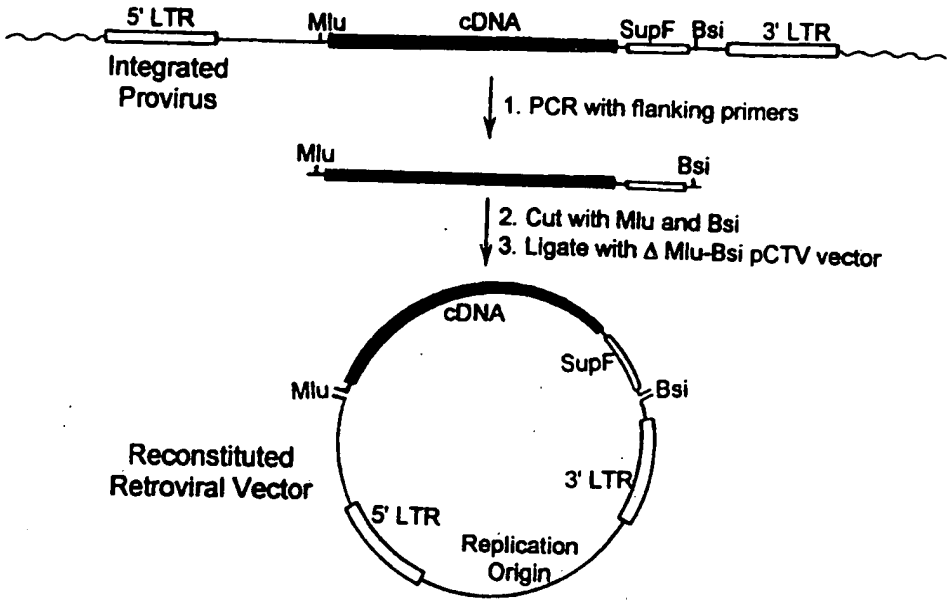


FIG. 3. Procedure for recovering cDNAs from integrated proviruses. Genomic DNA is prepared from infected cell clones and used as template for PCR amplification with a pair of primers complementary to vector sequences upstream of the site of cDNA insertion and downstream of the *supF* gene. After digestion with restriction enzymes *Mlu*I and *Bsi*WI, PCR products are gel purified and ligated between the same sites in pCTV3K, a variant of pCTV3 which has a kanamycin resistance gene rather than a *supF* gene between the *Mlu*I and *Bsi*WI sites. The formation of a complete plasmid which can be selected on the basis of suppressor tRNA activity is therefore dependent on the incorporation of the amplified *supF* gene, along with the accompanying cDNA insert. The plasmids isolated from transformed *E. coli* are fully functional retroviral vectors which can be immediately used for testing the activity of the recovered cDNA by transfection into packaging cells and infection of secondary recipients.

cific for the mRNA encoding the oncogenic middle T protein. This intron would be expected to be spliced out during viral transmission. Three recovered N-Ras cDNA clones and seven recovered polyomavirus cDNA clones were reconverted to retroviral form by transient transfection of GP+E-86 cells and then used to infect NIH 3T3 cells. All 10 of these clones were highly transforming, demonstrating that their oncogenic potential had been faithfully maintained through the process of transmission, PCR recovery, and retransmission.

Isolation of oncogenic cDNAs from retrovirally transferred libraries. Three separate cDNA libraries derived from hemopoietic cell lines were constructed in the pCTV vectors. An estimated 300,000 viral clones were generated from these libraries by transient transfection of GP+E-86 packaging cells and transferred to NIH 3T3 cells by infection. Cell cultures infected with these libraries developed a total of 83 distinctly transformed foci, with a variety of morphologies and growth rates, while uninfected cell cultures or those infected with viruses derived from the pCTV3 vector alone had 3- to 10-fold-lower frequencies of occurrence of transformed foci, equivalent to the spontaneous rate of transformation of the NIH 3T3 cells used in these experiments.

The foci of transformed cells which developed in the library-infected cultures were isolated and expanded, and transferred cDNAs within them were amplified by PCR and then cloned by ligation into the pCTV3K vector. Plasmids containing cDNA inserts of the expected sizes were tested for transforming activity by infection of NIH 3T3 cells, and the positive clones were sequenced to determine their identities. The mean sizes of recovered cDNAs with transforming activity were very similar to the mean sizes of the cDNAs in the originating libraries (Table 2). Therefore, there did not appear to be significant discrimination against the transmission of larger cDNAs under the competitive circumstances of library transfer.

Table 3 lists the 19 cDNAs with transforming activity which were recovered from PCRs performed on 47 of the transformed recipient cell clones. Twelve of the cDNAs encode proteins which have exact sequence identities to entries in the sequence databases. Three of these (Raf-1, Ect2, and Lck) are known to be able to oncogenically transform NIH 3T3 cells. Transforming activity of Raf-1 and Ect2 requires N-terminal truncation (28, 34, 47), as had occurred in the cDNAs cloned from our retroviral libraries. Lck transforms NIH 3T3 cells when its kinase activity is stimulated by mutations in negative regulatory sites (30), but the Lck cDNA which we cloned is predicted to encode a protein lacking kinase activity. Therefore, its weak transforming activity may occur through a different mechanism. The truncated Lck cDNA included three out-of-frame ATG codons upstream of the first ATG in the Lck-encoding reading frame. When these upstream ATGs were removed, which would be predicted to improve the translation efficiency of the cDNA (26), both the frequency and the potency of transformation were increased (Table 3). This finding implies that transformation was indeed mediated by the predicted translation product.

TABLE 2. Comparison of cDNA sizes in libraries versus sizes of recovered cDNAs

Library		Recovered cDNAs	
Name	Mean size (bp) of cDNAs	No.	Mean size (bp)
L16, L17	2,200	6	2,300
L18, L20	1,600	9	1,700
L19	2,200	2	1,600
L25	1,800	1	1,200
L27	2,500	1	2,400

TABLE 3. Transforming cDNAs isolated from retrovirally transferred cDNA libraries

cDNA[a]	Size (bp)	Identity	Transforming potency[b]	Transforming frequency[c]
TL16-8cC9	3,200	β-Catenin, N truncated	Moderate	0.05
TL16-8/ATG[d]		β-Catenin, ATG enhanced	Moderate	0.10
TL16-11cB1	1,600	No similarity to database sequences	Weak	0.01
TL16-13cB2	2,800	No similarity to database sequences	Moderate	0.01
TL16-25c2	1,900	Raf-1, N truncated	Strong	0.4
TL17-5cA3	2,500	Ect2, N truncated	Strong	0.5
TL17-11cB5	1,800	Lck, N truncated	Weak	0.005
TL17-11/ATG[d]		Lck, ATG enhanced	Moderate	0.03
TL18c2	800	Phospholipase C-γ_2, N + C truncated	Moderate	0.01
TL18c2/ATG[d]		Pholipase C-γ_2, ATG enhanced	Moderate	0.05
TL18-2c1	2,000	Thrombin receptor, full length	Strong	ND
TL18-3c1	1,600	Thrombin receptor, full length	Strong	0.1
TL18-5c1	1,700	Thrombin receptor, full length	Strong	ND
TL18-6cB4	1,200	Thrombin receptor, C truncated	Strong	ND
TL18-8c6	2,100	Thrombin receptor, full length	Strong	ND
TL18-9c1	1,800	CDC24 GNEF family, full length	Strong	0.2
TL18-10cA1	2,700	Low similarity to G_βs, antisense	Weak	0.02
TL19c39	1,300	Thrombin receptor, C truncated	Strong	ND
TL19-10c2	1,900	CDC24 GNEF homolog, N truncated	Strong	0.5
TL20-6cB8	1,500	Spi-2 protease inhibitor, antisense	Moderate	0.05
TL25-12-4-4	1,200	No similarity to database sequences	Strong	0.05
TL27-19-1-1	2,400	CDC24 GNEF homolog, C truncated	Strong	0.3

[a] TL16 and TL17 cDNAs were cloned from the GM979 erythroleukemic cell line cDNA library, TL18, TL19, and TL20 cDNAs were from the 32D myeloid cell line library, and the TL25 and TL27 cDNAs were from the B6SUtA$_1$ hemopoietic progenitor cell line library.

[b] Maximal density of cell growth and degree of morphological transformation observed in foci arising in NIH 3T3 cultures after infection with retroviral vectors carrying the indicated cDNA.

[c] Ratio of transformed colonies to total colonies arising in low-density NIH 3T3 cultures infected with pCTV3 retroviruses carrying the indicated cDNA. Colonies were counted and scored as transformed or nontransformed after 5 to 14 days under hygromycin selection. The spontaneous transformation frequencies of NIH 3T3 cultures infected with the pCTV3 vector alone ranged from 0.0001 to 0.001. The pCTV3/N-Ras virus had a transformation frequency of 0.7. ND, not determined.

[d] The sequences flanking presumptive start codons in these cDNAs were altered, as listed below, to optimize translation efficiency as follows: TL16-8cC9, CTGTTATGG; TL16-8/ATG, CCACCATGG, and one upstream ATG removed; TL17-11cB5, TGTGGATGG; TL17-11/ATG, CCACCATGG, and three upstream ATGs removed; TL18c2, TGGACATGC; TL18c2/ATG, CCACCATGG, and one upstream ATG removed. The cDNAs were also truncated immediately upstream of the modified start codons, to remove out-of-frame ATG codons which would be expected to reduce translation efficiency from the presumptive start codons.

The other cDNAs represented in the databases were derived from genes which were not previously identified as having transforming activity, and therefore their mechanisms of transformation are speculative. Thrombin receptor expression and its cleavage by thrombin in the serum used for cell culture presumably activates a G-protein-coupled mitogenic signalling pathway (27, 48) in NIH 3T3 cells, as has been demonstrated for ligand-stimulated serotonin and muscarinic receptors (18, 23). The transforming effects of the truncated β-catenin may be due to its disruption of the function of β-catenin-associated cell adhesion molecules such as E-cadherin (19), leading to transformation via loss of contact-induced growth repression. Alternatively, expression of truncated β-catenin may itself trigger a mitogenic signal transduction pathway in NIH 3T3 cells. The mechanism of action of the truncated phospholipase C-γ_2 is even more obscure, as it does not encode the catalytic domains required for mitogenic signal transduction (14). It could conceivably act by interfering with the function of an unknown inhibitor of the endogenous phospholipase C-γ_1 of NIH 3T3 cells. As seen for the Lck cDNA, the transforming activities of the truncated β-catenin and phospholipase C-γ_2 cDNAs were increased (Table 3) when upstream ATGs were removed and the sequence contexts of the presumptive initiation codons were optimized for translation (26).

The transforming activity of the antisense Spi-2 cDNA was presumably due to an inhibitory effect on the stability or translation of mRNAs encoding the Spi-2 protease inhibitor (20), leading to increased secreted protease activity and disruption of cell-cell or cell-matrix contacts which normally repress proliferation.

The remaining seven of the transforming cDNAs do not have high sequence identity with any entries in the DNA sequence databases. Three of these cDNAs encode proteins which include domains with homology to the CDC24 family of guanine nucleotide exchange factors (6) and therefore are presumed to transform NIH 3T3 cells by constitutively activating signal transduction through Ras-like proteins. One of the other novel clones has an antisense-oriented reading frame with weak but significant sequence similarity to beta subunits of G proteins and thus potentially acts as a suppressor of an endogenous G-protein-coupled signal transduction pathway which inhibits transformation of NIH 3T3 cells.

Of the 14 different types of cDNAs which were isolated in our screening of the retroviral cDNA libraries, 8 encoded proteins which are known or strongly implicated as positive inducers of mitogenic signals, and one (TL18-10cA1) may be an antisense suppressor of an inhibitor of mitogenic signals. This high yield of known or presumptive growth regulators confirms the validity of using the NIH 3T3 focus formation assay for detecting cDNAs encoding proteins with these functions. The fact that four of these cDNAs had not previously been cloned confirms our expectation that many novel growth regulators remain to be cloned via screening for transformation of NIH 3T3 cells or other transformation-susceptible cell lines. Given the proportion of growth regulators isolated in our screen, there is a high probability that the three novel cDNAs with no obvious coding sequence similarity to known proteins will also prove to be bona fide growth regulators. Alternatively, one or more of these cDNAs may encode or repress the expression of proteins which morphologically transform NIH 3T3 cells by

altering cell surface interactions (possible examples being the Spi-2 and β-catenin clones).

Some of the cDNAs described above appear to have acquired transforming activity via mutations arising in the process of cDNA library construction, i.e., truncation or inversion. The expression of such altered cDNAs can therefore be critical to the induction of a desired phenotype, either through the removal of negative regulatory domains from a phenotype inducer (e.g., truncated Raf-1 and Ect2) or by the generation of suppressors of phenotype repressors (e.g., antisense Spi-2). The possible dependence on specific mutation for the revelation of a phenotype such as oncogenic transformation means that only a small subset of the cDNAs derived from transcripts of a given gene may be functionally detectable. Furthermore, many of the cDNAs cloned in this study had transformation frequencies considerably less than 1. While this may in part be due to the transmission of a partially deleted virus or to promoter interference resulting in the inactivation of the LTR, in most cases it is likely due to critical thresholds of expression not being reached by most of the proviruses carrying the cDNA. Thus, the ability to transfer and screen very large cDNA libraries can be an essential requirement for the cloning of such cDNAs.

The NIH 3T3 cell line is obviously a very effective recipient cell for detecting and cloning oncogenes. This is largely because of its high susceptibility to transformation by the expression of proteins which activate the Ras-Raf signal transduction pathway. When expressed in the C3H10T1/2 fibroblast line, only the truncated Raf cDNA and the TL19-10c2 cDNA caused transformation, while all other cDNAs listed in Table 3 were negative. However, we have recently used the C127 epithelial cell line as a recipient for cDNA library transfer and have isolated one cDNA encoding a putative transcription factor which transforms this line but does not transform NIH 3T3 cells. This implies that the use of alternative recipient cells can either restrict or expand the spectrum of transforming cDNAs which can be detected and cloned.

The 19 transforming cDNAs listed in Table 3 were isolated from a relatively small number of transferred clones, representing just 7% of the 4 million clones in the available libraries. The scale of our library screens has been restricted not by the number of cDNA clones which can be transferred but rather by the number of transforming cDNAs which can be processed and analyzed. There are presumably many more transforming cDNAs remaining to be detected in these libraries. Of the 14 different types of transforming cDNAs isolated, multiple independent clones were obtained only in the case of the thrombin receptor, reflecting the relatively high abundance of its mRNA in 32D cells (25a) and the apparent ability of this receptor to transform NIH 3T3 cells without a requirement for specific truncating mutations.

Two major advantages were obtained through the use of retroviral library transfer for the isolation of cDNAs capable of inducing proliferation of NIH 3T3 cells. The first was the ability to screen very large numbers of cDNA clones on an equivalent number of recipient cells. This resulted in a high rate of cDNA-induced versus spontaneous transformation events and thus a high rate of recovery of transforming versus inert cDNAs. The second advantage was the relatively high levels of expression obtained with retrovirally transferred cDNAs, which in our hands are 10-fold greater on average than those obtained with the best available stable transfection vectors. Efficient expression can be critical for obtaining transformation and other targeted phenotypes in recipient cells, particularly when an antisense mechanism is involved.

A third major advantage of retroviral cDNA library transfer, not exploited in this study, is the potential to use cell lines which have previously been completely inaccessible to expression cloning because of low transfection efficiencies. There are only a limited number of fibroblast and epithelial cell lines which can be stably transfected with efficiencies of greater than 0.1% (10, 11). In contrast, almost all cell lines and some primary cell types can be infected with retroviruses. We have recently begun generating retroviral cDNA libraries with the BOSC23 packaging cell line, which was established for the specific purpose of yielding high viral titers via transient transfection (41). The viral titers of up to 10^6/ml obtained with these packaging cells are sufficiently high to enable the transfer of million-member cDNA libraries to hemopoietic cell lines and the detection of rare cDNA clones conferring selectable phenotypes such as conversion to growth in the absence of cytokines (50). Use of the BOSC23 cell line should permit the transfer of large cDNA libraries to cell lines representing many different lineages as well as to some primary cell types. This would make retroviral transfer of cDNA libraries uniquely suitable for the cloning of cDNAs which can impose new phenotypes or complement genetic defects in specialized cell types.

ACKNOWLEDGMENTS

We thank H. Land and S. Cory for providing the pBabeNeo and JZen retroviral vectors used in the assembly of the pCTV vectors, B. Seed for providing the *supF* gene and *BstXI* stuffer from his pCDM8 vector, M. Capecchi for the polyomavirus/TK promoter from pMC1, A. Bank and W. Pear for providing the GP+E-86 and BOSC23 packaging cell lines, J. Hassell for providing the polyomavirus early region clone, S. Nagata for the EF-1α promoter, and P. Lansdorp for CD24 detection reagents.

This work was funded by grants from the Medical Research Council of Canada, the National Cancer Institute of Canada, and the British Columbia Health Research Foundation and by a fellowship to I.W. from the Leukemia Research Foundation.

REFERENCES

1. Beck, E., G. Ludwig, E. A. Auerswald, B. Reiss, and H. Schaller. 1982. Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene 19:327–336.
2. Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller. 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. 61:1639–1646.
3. Bernard, H.-U., G. Krammer, and W. G. Rowekamp. 1985. Construction of a fusion gene that confers resistance against hygromycin B to mammalian cells in culture. Exp. Cell Res. 158:237–243.
4. Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1513–1523.
5. Bishop, J. M. 1991. Molecular themes in oncogenesis. Cell 64:235–248.
6. Boguski, M. S., and F. McCormick. 1993. Proteins regulating Ras and its relatives. Nature (London) 366:643–654.
7. Brown, B. A., R. W. Padgett, S. C. Hardies, C. A. Hutchinson, and M. H. Edgell. 1982. β-Globin transcript found in induced murine erythroleukemia cells is homologous to βh0 and βh1 genes. Proc. Natl. Acad. Sci. USA 79:2753–2757.
8. Chan, A. M.-L., M. Chedid, E. S. McGovern, E. S. Popescu, T. Miki, and S. A. Aaronson. Expression cDNA cloning of a serine kinase transforming gene. Oncogene 8:1329–1333.
9. Chan, A. M.-L., T. P. Fleming, E. S. McGovern, M. Chedid, T. Miki, and S. A. Aaronson. 1993. Expression cDNA cloning of a transforming gene encoding the wild-type Gα12 gene product. Mol. Cell. Biol. 13:762–768.
10. Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. 7:2745–2752.
11. Chen, C. A., and H. Okayama. 1988. Calcium phosphate-mediated gene transfer: A highly efficient transfection system for stably transforming cells with plasmid DNA. BioTechniques 6:632–638.
12. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162:156–159.
13. Cooper, C. S., M. Park, D. G. Blair, M. A. Tainsky, K. Huebner, C. M. Croce, and G. F. Vande Woude. 1984. Molecular cloning of a new transforming gene from a chemically transformed human cell line. Nature (London) 311:29–33.
14. Emori, Y., Y. Homma, H. Sorimachi, H. Kawasaki, O. Nakanishi, K. Suzuki,

and T. Takenawa. 1989. A second type of rat phosphoinositide-specific phospholipase C containing a src-related sequence not essential for phosphoinositide hydrolyzing activity. J. Biol. Chem. 264:21885–21890.
15. **Eva, A., and S. A. Aaronson.** 1985. Isolation of a new human oncogene from a diffuse B-cell lymphoma. Nature (London) 316:273–275.
16. **Gorman, C. M., B. H. Howard, and R. Reeves.** 1983. Expression of recombinant plasmids in mammalian cells is enhanced by sodium butyrate. Nucleic Acids Res. 11:7631–7648.
17. **Greenberger, J. S., M. A. Sakakeeny, R. K. Humphries, C. J. Eaves, and R. J. Eckner.** 1983. Demonstration of permanent factor-dependent multipotential (erythroid/neutrophil/basophil) hematopoietic progenitor cell lines. Proc. Natl. Acad. Sci. USA 80:2931–2935.
18. **Gutkind, J. S., E. A. Novotny, M. R. Brann, and K. C. Robbins.** 1991. Muscarinic acetylcholine receptor subtypes as agonist-dependent oncogenes. Proc. Natl. Acad. Sci. USA 88:4703–4707.
19. **Hamaguchi, M., N. Matsuyoshi, Y. Ohnishi, B. Gotoh, M. Takeichi, and Y. Nagai.** 1993. $p60^{v\text{-}src}$ causes tyrosine phosphorylation and inactivation on the N-cadherin-catenin cell adhesion system. EMBO J. 12:307–314.
20. **Inglis, J. D., M. Lee, D. R. Davidson, R. E. Hill.** 1991. Isolation of two cDNAs encoding novel alpha 1-antichymotrypsin-like proteins in a murine chondrocytic cell line. Gene 106:213–220.
21. **Jacobson, A.** 1987. Purification and fractionation of poly(A)$^+$ RNA. Methods Enzymol. 152:254–261.
22. **Johnson, G. R., T. J. Gonda, D. Metcalf, I. K. Hariharan, and S. Cory.** 1989. A lethal myeloproliferative syndrome in mice transplanted with bone marrow cells infected with a retrovirus expressing granulocyte-macrophage colony stimulating factor. EMBO J. 8:441–448.
23. **Julius, D., T. J. Livelli, T. M. Jessell, and R. Axel.** 1989. Ectopic expression of the serotonin 1c receptor and the triggering of malignant transformation. Science 244:1057–1062.
24. **Kay, R., and R. K. Humphries.** 1991. New vectors and procedures for isolating cDNAs encoding cell surface proteins by expression cloning in COS cells. Methods Mol. Cell. Biol. 2:254–265.
25. **Kay, R., P. M. Rosten, and R. K. Humphries.** 1991. CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. J. Immunol. 147:1412–1416.
25a. **Kirk, H., and R. Kay.** Unpublished data.
26. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266:19867–19870.
27. **LaMorte, V. J., E. D. Kennedy, L. R. Collins, D. Goldstein, A. T. Harootunian, J. H. Brown, and J. R. Feramisco.** 1993. A requirement for Ras protein function in thrombin-stimulated mitogenesis in astrocytoma cells. J. Biol. Chem. 268:19411–19415.
28. **Li, P., K. Wood, H. Mamon, W. Haser, and T. Roberts.** 1991. Raf-1: a kinase currently without a cause but not lacking in effects. Cell 64:479–482.
29. **Markowitz, D., S. Goff, and A. Bank.** 1988. A safe packaging line for gene transfer: separating viral genes on two different plasmids. J. Virol. 62:1120–1124.
30. **Marth, J. D., J. A. Cooper, C. S. King, S. F. Ziegler, D. Tinker, R. W. Overell, E. G. Krebs, and R. M. Perlmutter.** 1988. Neoplastic transformation induced by an activated lymphocyte-specific protein tyrosine kinase ($pp56^{lck}$). Mol. Cell Biol. 8:540–550.
31. **Martin-Zanca, D., S. H. Hughes, and M. Barbacid.** 1986. A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. Nature (London) 319:743–748.
32. **Miki, T., T. P. Fleming, D. P. Bottaro, J. S. Rubin, D. Ron, and S. A. Aaronson.** 1991. Expression cDNA cloning of the KGF receptor by creation of a transforming autocrine loop. Science 251:72–75.
33. **Miki, T., T. P. Fleming, M. Crescenzi, C. J. Molloy, S. B. Blam, S. H. Reynolds, and S. A. Aaronson.** 1991. Development of a highly efficient expression cDNA cloning system: application to oncogene isolation. Proc. Natl. Acad. Sci. USA 88:5167–5171.
34. **Miki, T., C. L. Smith, J. E. Long, A. Eva, and T. P. Fleming.** 1993. Oncogene *ect2* is related to regulators of small GTP-binding proteins. Nature (London) 362:462–465.
35. **Miller, A. D., and G. J. Rosman.** 1989. Improved retroviral vectors for gene transfer and expression. BioTechniques 7:980–990.
36. **Miyoshi, J., T. Higashi, H. Mukai, T. Ohuchi, and T. Kakunaga.** 1991. Structure and transforming potential of the human *cot* oncogene encoding a putative protein kinase. Mol. Cell. Biol. 11:4088–4096.
37. **Mizushima, S., and S. Nagata.** 1990. pEF-BOS, a powerful mammalian expression vector. Nucleic Acids Res. 18:5322.
38. **Morgenstern, J. P., and H. Land.** 1990. Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. Nucleic Acids Res. 18:3587–3596.
39. **Murphy, A. J. M., and A. Efstratiadis.** 1987. Cloning vectors for expression of cDNA libraries in mammalian cells. Proc. Natl. Acad. Sci. USA 84:8277–8281.
40. **O'Bryan, J. P., R. A. Frye, P. C. Cogswell, A. Neubauer, B. Kitch, C. Prokop, R. Espinosa III, M. M. LeBeau, H. S. Earp, and E. T. Liu.** 1991. *axl*, a transforming gene isolated from primary human myeloid leukemia cells, encodes a novel receptor tyrosine kinase. Mol. Cell. Biol. 11:5016–5031.
41. **Pear, W. S., G. P. Nolan, M. L. Scott, and D. Baltimore.** 1993. Production of high-titer helper-free retroviruses by transient transfection. Proc. Natl. Acad. Sci. USA 90:8392–8396.
42. **Rayner, J. R., and T. J. Gonda.** 1994. A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. Mol. Cell. Biol. 14:880–887.
43. **Seed, B.** 1983. Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo. Nucleic Acids Res. 11:2427–2445.
44. **Seed, B.** 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. Nature (London) 329:840–842.
45. **Seed, B., and A. Aruffo.** 1987. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. Proc. Natl. Acad. Sci. USA 84:3365–3369.
46. **Shimizu, K., M. Goldfarb, M. Perucho, and M. Wigler.** 1983. Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. Proc. Natl. Acad. Sci. USA 80:383–387.
47. **Stanton, V. P., Jr., D. W. Nichols, A. P. Laudano, and G. M. Cooper.** 1989. Definition of the human *raf* amino-terminal regulatory region by deletion mutagenesis. Mol. Cell. Biol. 9:639–647.
48. **Thien-Khai, H., D. T. Hung, V. I. Wheaton, and S. R. Coughlin.** 1991. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. Cell 64:1057–1068.
49. **Thomas, K. R., and M. R. Capecchi.** 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51:503–512.
50. **Tognon, C., and R. Kay.** Unpublished data.
51. **Vogeli, G., E. Horn, M. Laurent, and P. Nath.** 1985. Recombinant DNA techniques: storage and screening of cDNA libraries with large numbers of individual colonies from initial transformations. Anal. Biochem. 151:442–444.

Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 11509–11513, December 1995
Biochemistry

Isolation of virus-neutralizing RNAs from a large pool of random sequences

(*in vitro* selection/nuclease-resistant RNA analogs/antiviral agents)

WEIHUA PAN*, REBECCA C. CRAVEN†, QIU QIU*‡, CAROL B. WILSON†, JOHN W. WILLS†, SERGUEI GOLOVINE*, AND JIN-FENG WANG*§

Departments of *Biochemistry and Molecular Biology and †Microbiology and Immunology, Pennsylvania State University College of Medicine, Hershey, PA 17033

Communicated by Olke C. Uhlenbeck, University of Colorado, Boulder, CO, July 27, 1995 (received for review April 27, 1995)

ABSTRACT **RNA and ribonuclease-resistant RNA analogs that bound and neutralized Rous sarcoma virus (RSV) were isolated from a large pool of random sequences by multiple cycles of *in vitro* selection using infectious viral particles. The selected RNA pool of RSV-binding sequences at a concentration of 0.16 μM completely neutralized the virus. Of 19 sequences cloned from the selected pool, 5 inhibited RSV infection. The selected RNA and RNA analogs were shown to neutralize RSV by interacting with the virus, rather than by adversely affecting the host cells. The selection of the anti-RSV RNA and RNA analogs by intact virions immediately suggests the potential application of this approach to develop RNA and RNA analogs as inhibitors of other viruses such as human immunodeficiency virus.**

In comparison to the large numbers of antimicrobial agents against bacteria, few effective antiviral drugs have been developed (1, 2). The most widely used approach to the discovery of antiviral agents has been the empirical screening of chemically diverse classes of synthetic and natural compounds for the ability to inhibit virus replication (2, 3). However, the empirical screening is usually tedious and inefficient. Other approaches such as rational drug design based on the known structures of viral and cellular proteins are promising but to date have not been largely successful (3).

RNA molecules have a remarkable diversity of structure and function (4, 5). The structural complexity of a combinatorial sequence library of single-stranded RNA oligonucleotides provides the potential to select for molecules that bind defined targets with high affinity and specificity (6–8). Because the selected RNA molecules can be amplified and subjected to further selection, multiple cycles of such selection and amplification can lead to the isolation of the best target-binding RNA molecules from the library. This *in vitro* selection–amplification procedure has been previously used to isolate RNA (and DNA) molecules that bind a variety of small molecules and purified proteins (9–12). We have extended the procedure to a new type of target, an intact virus, to isolate the RNA molecules and the nuclease-resistant RNA analogs that specifically bind and neutralize the virus.

Rous sarcoma virus (RSV), an avian retrovirus, is among the best-studied members of the family Retroviridae, which includes human immunodeficiency virus (HIV). All members of this family are enveloped viruses that display on their surface glycoproteins which are required for binding to a specific receptor on a susceptible cell and for fusion with the cellular membrane (13, 14). Here we use RSV to demonstrate that without knowledge of the structures of viral proteins, antiviral RNAs and RNA analogs can be isolated systematically and effectively from a large pool of random sequences, first by using intact viral particles to select from the sequence pool the RNAs and RNA analogs that bind specifically to the virus, and subsequently by screening the selected molecules for the ability to neutralize the virus. We reason that by binding to virus, some of the selected RNAs and RNA analogs may change the structures of viral surface proteins so that these proteins can no longer function in steps critical for viral infection, such as viral attachment and virus–cell membrane fusion. Alternatively, some of the structural changes may trigger pathways to inhibit the steps which normally occur after virus internalization, such as the uncoating and the expression of the virus genome. Both mechanisms have been previously suggested for many similar virus neutralization reactions induced by the binding of antibodies (15, 16).

MATERIALS AND METHODS

Materials. Nucleoside 5′-triphosphates were purchased from United States Biochemical. T7 RNA polymerase was purified from the overproducing *Escherichia coli* strain BL21/pAR1219 (17). SuperScript II reverse transcriptase was from GIBCO/BRL. *Taq* DNA polymerase was from Promega. Human plasma was provided by David Farrell (Pennsylvania State University). Prague A strain of RSV and anti-RSV serum were prepared as described (18–20).

Construction of a Large RNA Pool of Random Sequences. A DNA library containing ~5×10^{16} sequences (2.5 mg of DNA) was constructed by automated solid-state synthesis. The sequence diversity was generated by randomizing a central 40-nt region (40N) of the 87-nt oligomer 5′-GCC-GGA-TCC-GGG-CCT-CAT-GTC-GAA-40N-TTG-AGC-GTT-TAT-TCT-GAG-CTC-CC. The 5′-end primer (5′-GCC-GGA-TCC-GGG-CCT-CAT-GTC-GAA-3′) and 3′-end primer (5′-CCG-AAG-CTT-AAT-ACG-ACT-CAC-TAT-AGG-GAG-CTC-AGA-ATA-AAC-GCT-CAA-3′, containing a T7 RNA polymerase promoter) were made for the polymerase chain reaction (PCR) to amplify the 87-mer. Large-scale PCR of the DNA library resulted in double-stranded DNA (~5×10^{15} sequences) which was transcribed by T7 RNA polymerase to generate a pool of multiple copies of ~10^{15} RNA sequences (random pool).

Transcription of RNAs and 2′-Fluoropyrimidine-Containing RNAs (2′-F-RNAs). Unmodified RNAs were transcribed *in vitro* by T7 RNA polymerase at 37°C for 2–3 hr in a buffer containing 1 mM each ATP, CTP, GTP, and UTP; 60 mM Tris·HCl (pH 8.0), 12 mM $MgCl_2$, 1 mM spermidine, 5 mM dithiothreitol, and 0.001% Triton X-100. 2′-F-RNAs were transcribed at 35°C for 8 hr in the same buffer except that 1

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RSV, Rous sarcoma virus; HIV, human immunodeficiency virus; 2′-F-RNA, 2′-fluoropyrimidine-containing RNA.

‡Present address: Residency Program, Department of Medicine, Polyclinic Medical Center, Harrisburg, PA 17110.
§To whom reprint requests should be addressed.

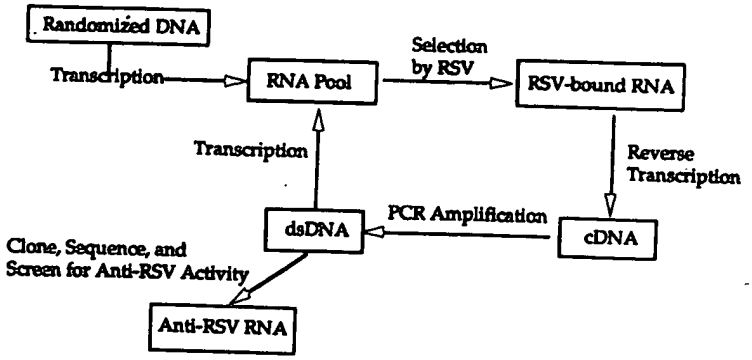


FIG. 1. *In vitro* selection–amplification procedure for isolating RSV-binding and -neutralizing RNAs and RNA analogs from a combinatorial sequence library. dsDNA, double-stranded DNA.

mM 2′-F-CTP and 2′-F-UTP replaced CTP and UTP. The full-length transcripts were purified by polyacrylamide gel electrophoresis.

***In Vitro* RNA Selection by RSV Particles.** The RNA (or 2′-F-RNA) sequence pool was incubated with RSV (Prague A strain) for 10 min at 37°C in 2.5 mM $MgCl_2$/100 mM NaCl/20 mM Tris·HCl, pH 7.5. The mixture was then filtered through a prewetted nitrocellulose membrane. The RNA–virus complexes were recovered from the nitrocellulose membrane after the unbound RNAs were washed away. The RNA molecules which bound anything else but RSV (e.g., the nitrocellulose membrane, the cellular proteins contaminating the virus solution) were removed from the RNA pool by a background selection either prior to or after the selection by virus. The RNA–RSV complexes were denatured and viral proteins were removed by phenol/chloroform extraction. The RNAs were then recovered by ethanol precipitation and reverse-transcribed to their cDNAs by SuperScript II reverse transcriptase. The cDNAs were PCR-amplified and subsequently transcribed to give a new pool of RNA which was enriched for better RSV-binding sequences. This new pool of RNA was used for selection by intact RSV to begin the next cycle. The selection stringency was increased by lowering the RSV concentration as the number of selection cycles increased. The RNA–virus complexes were also separated from the unbound RNA by high-speed centrifugation rather than filtering during cycles 4 and 9; the RNA–virus complexes formed a pellet along with any unbound virus while the unbound RNA remained in the supernatant. After 12 cycles, the selected RNA pool was tested for the ability to bind and neutralize RSV.

Binding of the Selected RNAs to RSV. The ^{32}P-labeled RNA pool (or a cloned RNA sequence) obtained after 12 cycles of selection was incubated at 0.2–0.5 nM with various concentrations of RSV for 15 min at 37°C in 2.5 mM $MgCl_2$/100 mM NaCl/20 mM Tris·HCl, pH 7.5. The RNA–virus mixtures were filtered on prewetted nitrocellulose membranes; >95 of the RNA–RSV complexes were retained on the membrane, and the unbound RNAs were washed away by the binding buffer. The radioactivity retained on the membranes was measured by liquid scintillation counting and corrected for any background due to the binding of RNAs to nitrocellulose membrane.

Virus-Neutralizing Activity of the Selected RNAs. Fresh RSV suspension (Prague A strain, $\sim 10^5$ focus-forming units/ml) was supplemented with 2.5 mM $MgCl_2$, mixed with the selected RNA pool or a specific RNA sequence, and incubated at 37°C for 15 min. The RNA–virus mixture was placed onto 10^6 quail fibrosarcoma cells (QT6 cell line), incubated for 1 hr at 37°C in a 5% CO_2 atmosphere, and then replaced with complete growth medium (19). At 18 hr postinfection, the cells were labeled with L-[^{35}S]methionine for 2 hr. Viral proteins in cell lysates and media were recovered by immunoprecipitation with an anti-RSV serum and analyzed on 12% polyacrylamide/0.1% SDS gels (19, 20). The yields of viral proteins on the gels were estimated by densitometry.

RESULTS

***In Vitro* Selection of RSV-Binding RNA Sequences.** An RNA pool of $\sim 10^{15}$ different sequences, each 87 nt long, was subjected to selection by sucrose-gradient purified RSV particles to isolate the RNA molecules that bound the virus. The RSV-binding RNA sequences were then amplified and subjected to the next cycle of selection at a lower RSV concentration to increase the selection stringency (Fig. 1). Multiple cycles of selection and amplification resulted in an exponential increase of the best RSV-binding RNA molecules. After 12 cycles, the RSV concentration required to bind 50% of ≈0.5 nM selected RNA pool (dissociation constant, K_d) decreased to ≈3 μg of viral protein per ml, which was at least 1000 times lower than that required to bind 50% of the pool of random sequences (Fig. 2). The RSV concentration was expressed in terms of total viral protein concentration because the molar number of viral particles could not be accurately measured. Sequencing of the selected RNA pool revealed that the RNA sequences were no longer random and particular sequences appeared to dominate the population. The selected RNA pool was then tested for its ability to neutralize RSV.

Neutralization of RSV by the Selected RNA Pool. In the absence of any added RNA, avian cells (QT6) were efficiently infected by RSV as shown by the large amount of newly

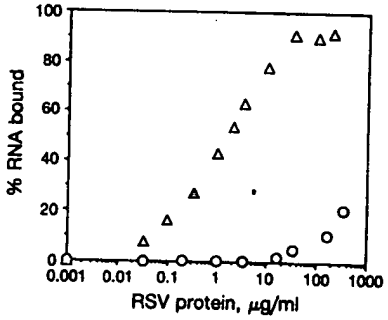


FIG. 2. Binding of 0.2–0.5 nM ^{32}P-labeled RNA to RSV. △, Selected RNA pool obtained after 12 selection–amplification cycles; ○, random RNA pool. Because the detection limit for the ^{32}P-labeled RNA was about 0.05 nM and the affinity between the selected RNA pool and RSV was high, the measured percentage of the bound selected RNA at low concentrations of RSV may be limited by the available binding sites on the virus rather than by the dissociation constant (K_d). Thus, the K_d of 3 μg of viral protein per ml derived from such measurements may be an estimate of its upper limit.

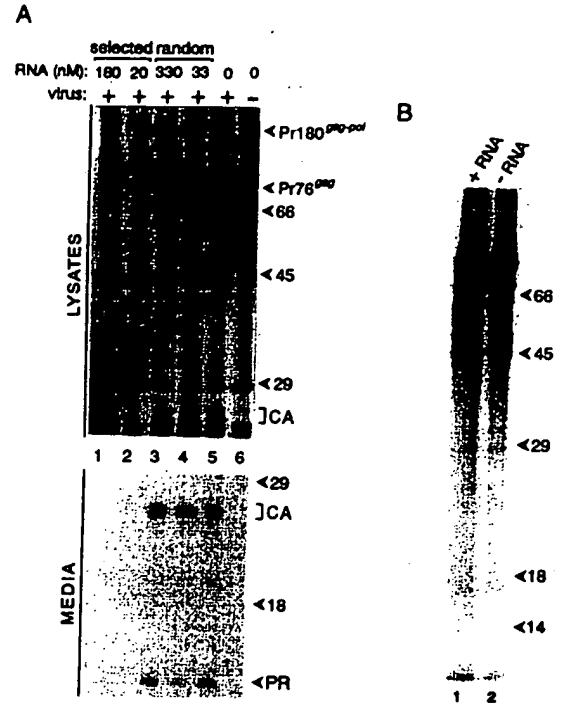


FIG. 3. Effect of the selected RNA pool on RSV infection of QT6 quail cells and on host cell protein synthesis. (*A*) Inhibition of RSV production. Viral proteins in cell lysates and media were recovered by immunoprecipitation with an anti-RSV serum and analyzed on 12% polyacrylamide/0.1% SDS gels. $Pr76^{gag}$, full-length Gag protein; $Pr180^{gag\text{-}pol}$, Gag-Pol fusion protein; CA (capsid proteins) and PR (protease), mature cleavage products derived from $Pr76^{gag}$. The RNA concentration in each virus sample is shown above each lane. However, the effective RNA concentration in each sample should be lower than the concentration shown here and in Fig. 4*A*, because up to 70% of the initially added RNA was found to be degraded within 75 min by nucleases in the RSV suspension. Unlike the sucrose gradient-purified RSV used for the selection, the RSV suspension used for the infection assay was freshly prepared and not purified, and thereby was more likely to be contaminated with ribonucleases from cell culture. (*B*) Effect of the selected RNA pool on the host cell protein synthesis. QT6 cells were treated in the absence of RSV with (lane 1) and without (lane 2) 500 nM selected RNA pool under the same conditions as used in *A*. After radiolabeling, total cellular proteins were analyzed by electrophoresis and autoradiography.

synthesized viral Gag ($Pr76^{gag}$) and Gag-Pol ($Pr180^{gag\text{-}pol}$) proteins in cell lysates, and capsid proteins (CA) and viral protease (PR) of the Gag cleavage products in growth media (Fig. 3*A*, lane 5) (19, 20). The presence of random-sequence RNA molecules at 330 nM had no effect, within experimental fluctuation, on the production of either the viral proteins or the background cellular proteins such as the 29-kDa band (Fig. 3*A*, lanes 3 and 4). In contrast, the selected RNA pool at 20 nM reduced the yield of viral proteins by 85–92% (Fig. 3*A*, lane 2, estimated by densitometry). At 160 nM, the selected RNA pool completely blocked RSV production in this assay (Fig. 3*A*, lane 1). The selected RNA pool did not affect the yield of the cellular proteins to a detectable extent (Fig. 3*A*, compare the background bands in lanes 1 and 2 with those in lane 5). Thus the RNA pool of selected sequences appeared to block RSV production by neutralizing the virus, rather than by adversely affecting the host cells. Such specific inhibition of RSV infection by the selected RNA pool and the cloned individual RNA sequences described below was reproduced in five and four independent sets of experiments, respectively.

The effect of the selected RNA pool on the host cell's growth and susceptibility to viral infection was further tested.

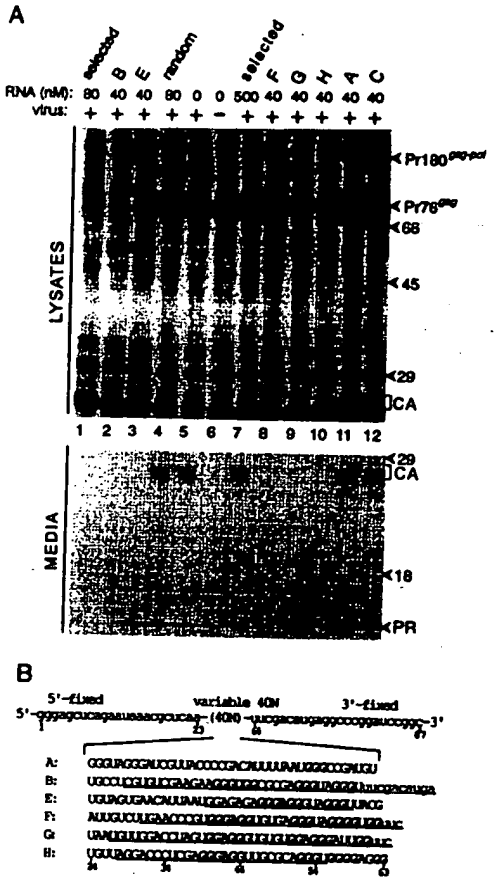


FIG. 4. (*A*) Inhibition of RSV by individual RNA sequences. The infections were performed in the same way as for the selected RNA pool (Fig. 3). In this particular experiment, viral protein $Pr76^{gag}$ overlapped with some of the background cellular proteins (comparison with lane 6). For lane 7, the QT6 cells were treated with the selected RNA pool (500 nM) for 1 hr and washed with medium prior to infection by RSV. (*B*) Nucleotide sequences of the RNAs that exhibited neutralizing activity (B, E, F, G, and H). Sequence A bound both RSV and nitrocellulose membrane and was the most abundant sequence in the selected pool but did not inhibit RSV. Each sequence consists of a 5′ fixed region, the variable 40 nt (40N) region (uppercase letters), and a 3′ fixed region. Bold letters show the groups of guanine nucleotides that appear at similar positions in at least two sequences. The underlined regions are the minimum sequences required for specific binding to RSV, as determined by the partial alkaline hydrolysis method (21).

Exposure of uninfected QT6 cells to the selected RNA pool at 500 nM did not affect cellular protein synthesis (Fig. 3*B*). Furthermore, pretreatment of cells with the selected RNA pool at 500 nM for 1 hr prior to infection by RSV did not change the yield of viral proteins (Fig. 4*A*, compare lane 7 with lane 5), indicating that the selected RNA pool had no effect on the susceptibility of the cells to RSV infection. In addition, the selected RNA pool showed no higher affinity for the host cells than did the random RNA pool (unpublished data), indicating no specific binding to the cells. All these results were consistent with the conclusion that the antiviral activity of the selected RNAs was due to their specific interactions with RSV.

Specific RNA Sequences That Neutralize RSV. To identify specific RSV-neutralizing sequences in the selected RNA pool, we cloned the DNA templates for the selected RNA pool into pUC19 vectors and sequenced 36 clones. The most prevalent sequence (sequence A) appeared in 9 clones, followed by

sequences B, C, D, and E, which were found in 4, 4, 3, and 2 clones, respectively. Each of the other 14 clones had a unique sequence. Among the 19 sequences identified from the 36 clones, RNA transcripts of 5 sequences (B, E, F, G, and H) showed RSV-neutralizing activity (Fig. 4). Because in the particular experiment shown in Fig. 4*A* viral protein $Pr76^{gag}$ overlapped with some of the background cellular proteins, which vary from one experiment to another (lane 6), the comparison of RSV production between different lanes may be better judged by the yields of $Pr180^{gag\text{-}pol}$ (in the lysates) and capsid proteins and protease (in the media). It was clear that sequences E, F, and G at 40 nM inhibited RSV as effectively as the selected RNA pool at 80 nM (Fig. 4*A*, compare lanes 3, 8, and 9 with lane 1). Side-by-side experiments showed that sequences B and H were 2- to 3-fold as effective as sequences

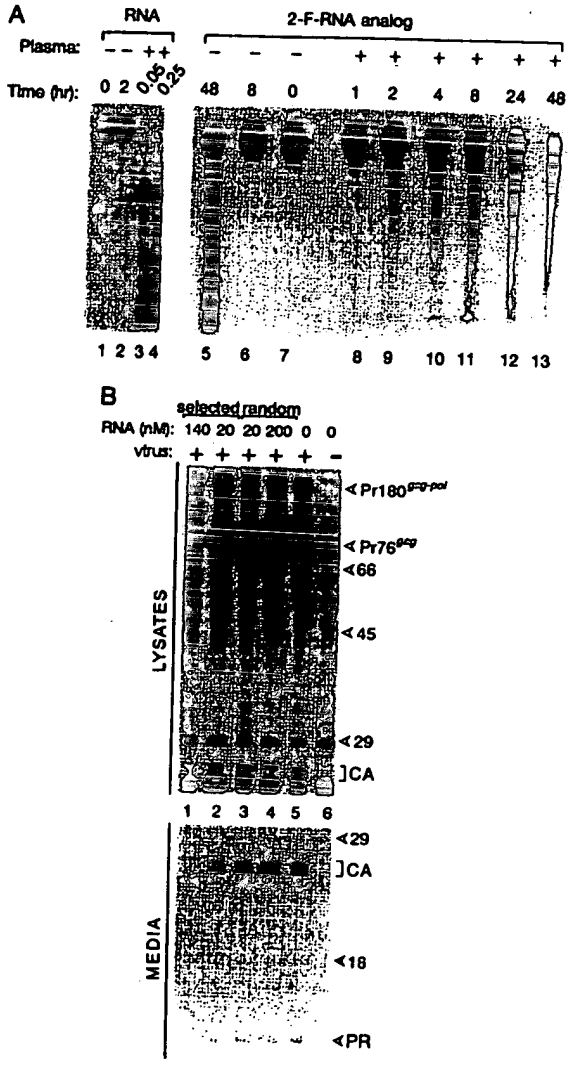


FIG. 5. (*A*) Stability of RNA and 2′-F-RNA of sequence H against degradation by nucleases in human plasma. The RNA and the RNA analog were incubated in human plasma supplemented with 2.5 mM $MgCl_2$ at 37°C for the times shown. Lane 5 shows the Mg^{2+}-enhanced background hydrolysis in 48 hr. The half-life of the RNA analog in human plasma (22 ± 4 hr) was about the same as those of the analog pools of the selected and random sequences. (*B*) Inhibition of RSV by the analog pool of the selected sequences. The analog pool was transcribed from the pool of sequences that was generated after 12 cycles of selection with an unmodified RNA pool.

E, F, and G in neutralizing the virus (unpublished data). Sequence A bound to RSV with an affinity comparable to that of sequences B and H ($K_d \sim 2$–3 μg of viral protein per ml) but had at least 4 times higher affinity for nitrocellulose membrane, which might explain why it was the most frequent sequence among the clones. Comparison of the cl ned sequences revealed that groups f guanine nucleotides appeared at similar positions in all the RSV-neutralizing RNA sequences (Fig. 4*B*), but no consensus could be discovered among their secondary structure (base-pairing) diagrams that were generated by energy minimization (22).

Neutralization of RSV by 2′-F-RNA Analogs. One of the potential problems in developing RNA as a therapeutic agent is the rapid degradation of unmodified RNA by ribonucleases present *in vivo*. This problem may be solved by a number of approaches, such as incorporating modified nucleotides into the RNA chain (23, 24). Indeed, the incorporation of 2′-fluoro-2′-deoxycytidine and 2′-fluoro-2′-deoxyuridine into the RNA chain by T7 transcription increased the RNA stability against the nuclease digestion by more than 2 orders of magnitude (Fig. 5*A*). A similar increase in RNA stability was observed when either the RNA pool of random sequences or the selected RNA pool or sequence H was tested (unpublished data), indicating little dependence on RNA sequence (presumably each RNA chain contained about the same number of pyrimidines).

A pool of 2′-F-RNAs of the same selected sequences as used in Fig. 3*A* inhibited RSV infection by ~85% (comparison of lane 1 with lanes 3, 4, and 5 of Fig. 5*B* by densitometry) and ~75% (unpublished data) when used at 140 nM and 100 nM, respectively. In contrast, the level of 85% inhibition was achieved by only 20 nM unmodified RNA pool (Fig. 3*A*, lane 2), indicating that the analog pool was 7–10 times less effective in neutralizing RSV. The RNA analogs of random sequences at 200 nM had no effect on RSV infection (Fig. 5*B*, lanes 3 and 4). Neither selected nor random sequences in the analog pool at up to 200 nM affected host cell protein synthesis (Fig. 5*B*, lanes 1 and 4). Because each of the 12 cycles of selection was performed with unmodified RNAs, it was likely that the lower anti-RSV activity of the RNA analogs was due to the RNA structural changes induced by the 2′-fluoro modification in pyrimidines. Nevertheless, the neutralizing activity of the analogs at high concentrations implied that some structural features of the unmodified RNA important for antiviral activity remained.

To test whether selection of RNA analogs would lead to a new pool with improved neutralizing activity, we transcribed a pool of 2′-F-RNAs from the sequences that were generated after 9 cycles of selection by RSV from an unmodified RNA pool, and then continued the selection with the analog pool for another 3 cycles. The resulting pool of 2′-F-RNAs inhibited RSV production by 75% when used at 22 nM (unpublished data), whereas the same level of inhibition was achieved at 100 nM for the analog pool transcribed after 12 cycles of selection with an unmodified RNA pool. Thus, the selection with 2′-F-RNAs for only 3 cycles improved the anti-RSV activity by 4.5-fold. An even more potent anti-RSV analog pool would have likely been obtained if all the 12 cycles of selection had been done with analog pools. Nevertheless, it is clear that ribonuclease-resistant antiviral RNA analogs can be obtained either by selecting from an unmodified RNA pool and then transcribing the selected sequences with modified nucleotides or by direct selection from an RNA analog pool.

DISCUSSION

Because of the permeation problem imposed by the RSV membrane, the selected RNAs are expected to inhibit RSV infection by binding to the virus surface and not by entering the virus. However, the identification of the RNA binding sites on

RSV and f the roles of these viral components in the virus infection has not yet been completed.

Selection for RNAs Inhibiting Multiple Strains of Virus. Although the selections were perf rmed with the Prague A strain of RSV, preliminary results indicated that the selected RNAs also inhibited the Prague C strain of RSV. The RNA concentration required t inhibit the Prague C strain was about 15–20 times higher than that needed to achieve the same level of inhibition of the Prague A strain (unpublished data). The surface glycoprotein (gp85) of these two strains contains conserved regions (~95% homology) as well as three or four small variable regions (subgroup-determining regions) (25, 26). If the selected RNAs and analogs neutralize RSV by interacting with the surface protein, it is likely that a small number (e.g., ~5%) of the Prague A-neutralizing sequences in the selected pool bind to the conserved regions of gp85 and thus also neutralize the Prague C strain.

The selection may also be performed by alternating between several strains of a virus during the different cycles of selection to specifically isolate the RNAs and RNA analogs that would bind to the viral conserved structural motifs and thereby inhibit diverse strains of the virus (S. Dallabrida, J. C. Sanders, E. M. Eyster, and J.-F.W., unpublished data). The isolation of RNAs and RNA analogs that are effective against diverse viral strains will be particularly necessary for inhibiting viruses such as HIV, which has a high mutation rate and readily develops drug-resistant strains. Such selection is possible and practical because a very large structurally diverse library is used and the selection is a rapid process.

Selection by Intact Biological Entities. Although an isolated viral protein can be used in the *in vitro* selection for the protein-binding RNAs and some of the selected sequences may inhibit viral infection, there are at least three advantages to the use of intact viral particles in the selection. (*i*) The selection by intact virus does not require a full understanding of the usually very complex mechanism of viral infection, whereas the selection by an isolated viral protein is limited to only a few cases in which the proteins responsible for the viral infection have been identified and isolated. (*ii*) Because of structural differences in the protein (27, 28), the RNA selected to bind a viral protein in its purified form may not interact with the protein complexed on a virion. (*iii*) The selection using intact virus may lead to the identification of viral components that have not previously been known for their critical roles in viral infection. The success of isolating numerous anti-RSV RNAs and RNA analogs from a large pool of random sequences immediately suggests the potential application of this approach to isolate RNA and nuclease-resistant RNA analogs against other viruses such as HIV.

The selection of RNAs (or DNAs) and their analogs may also be performed by other biological entities such as bacteria, yeast, and tumor cells (S.G., W.P., and J.-F.W., unpublished data). The selected nucleic acids and their analogs would bind to these cells and may change their biological functions by affecting cellular interactions such as signal transduction pathways. Therefore, the selection of RNAs and RNA analogs by intact viruses or cells may provide valuable tools for studying virus–cell and intercellular interacti ns. The selected RNAs and analogs may be used to control such interacti ns so that the infecti n by virus, bacterium, r yeast or the proliferation or death of cells may be inhibited.

We thank J. E. Hopper, D. Spector, M. Katzman, and O. Uhlenbeck for discussions and A. K. Hopper, R. Shiman, and L. Parent for comments on the manuscript. This work was supported by Pennsylvania State University Start-up Funds (J.-F.W.), National Institutes of Health Grant CA47482 (J.W.W.), American Cancer Society Grant FRA427 (J.W.W.), and funds from Pennsylvania State University College of Medicine (R.C.C.).

1. Fields, B. N. (1990) *Fields Virology* (Raven, New York).
2. van der Sijs, I. H. & Wiltink, E. H. (1994) *Int. J. Biochem.* **26**, 621–630.
3. Johnston, M. I. & Hoth, D. F. (1993) *Science* **260**, 1286–1293.
4. Gesteland, R. F. & Atkins, J. F., eds. (1993) *The RNA World* (Cold Spring Harbor Lab. Press, Plainview, NY).
5. Illangasekare, M., Sanchez, G., Nickles, T. & Yarus, M. (1995) *Science* **267**, 643–647.
6. Kenan, D. J., Tsai, D. E. & Keene, J. D. (1994) *Trends Biochem. Sci.* **19**, 57–64.
7. Irvine, D., Tuerk, C. & Gold, L. (1991) *J. Mol. Biol.* **222**, 739–761.
8. Joyce, G. F. (1989) *Gene* **82**, 83–87.
9. Kinzler, K. W. & Vogelstein, B. (1989) *Nucleic Acids Res.* **17**, 3645–3653.
10. Ellington, A. D. & Szostak, J. W. (1990) *Nature (London)* **346**, 818–822.
11. Tuerk, C. & Gold, L. (1990) *Science* **249**, 505–510.
12. Prudent, J. R., Uno, T. & Schultz, P. G. (1994) *Science* **264**, 1924–1927.
13. Varmus, H. (1988) *Science* **240**, 1427–1435.
14. Schupbach, J. (1990) *Human Retrovirology: Facts and Concepts* (Springer, Berlin).
15. Dimmock, N. J. (1992) *Curr. Top. Microbiol. Immunol.* **183**, 1–149.
16. Levy, J. A. (1994) *HIV and the Pathogenesis of AIDS* (Am. Soc. Microbiol., Washington, DC), pp. 153–158.
17. Davanloo, P., Rosenburg, A. H., Dunn, J. J. & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039.
18. Smith, R. E. (1979) *Methods Enzymol.* **58**, 393–403.
19. Craven, R. C., Leure-DuPree, A. E., Erdie, C. R., Wilson, C. B. & Wills, J. W. (1993) *J. Virol.* **67**, 6246–6252.
20. Wills, J. W., Craven, R. C. & Achacoso, J. A. (1989) *J. Virol.* **63**, 4331–4343.
21. Pan, T. & Uhlenbeck, O. C. (1992) *Biochemistry* **31**, 3887–3895.
22. Zuker, M. (1989) *Science* **244**, 48–52.
23. Pieken, W. A., Olsen, D. B., Benseler, F., Aurup, H. & Eckstein, F. (1991) *Science* **253**, 314–317.
24. Lin, Y., Qiu, Q., Gill, S. C. & Jayasena, S. D. (1994) *Nucleic Acids Res.* **22**, 5229–5234.
25. Dorner, A. J. & Coffin, J. M. (1986) *Cell* **45**, 365–374.
26. Bova, C. A., Olsen, J. C. & Swanstrom, R. (1988) *J. Virol.* **62**, 75–83.
27. Moore, J. P., McKeating, J. A., Huang, Y., Ashkenazi, A. & Ho, D. D. (1992) *J. Virol.* **66**, 235–243.
28. Orloff, S. L., Kennedy, M. S., Belperron, A. A., Maddon, P. J. & McDougal, J. S. (1993) *J. Virol.* **67**, 1461–1471.

Cell, Vol. 33, 153–159, May 1983, Copyright 1983 by MIT

0092-8674/83/0153-07 $02.00/0

Construction of a Retrovirus Packaging Mutant and Its Use to Produce Helper-Free Defective Retrovirus

Richard Mann*†, Richard C. Mulligan†, and David Baltimore*†

* Whitehead Institute for Biomedical Research
Cambridge, Massachusetts 02139

† Center for Cancer Research
and Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

A mutant of Moloney murine leukemia virus (M-MuLV), pMOV-ψ^-, was constructed by deletion of about 350 nucleotides from an infectious proviral DNA clone between the putative *env* mRNA 5′ splice site and the AUG that initiates the coding sequence for Pr65^{gag}. Although the parent wild-type proviral clone, pMOV-ψ^+, quickly causes a high level of reverse-transcriptase-containing virus particles to be released from transfected NIH/3T3 cells, transfection of pMOV-ψ^- into these cells initially results in very little release. By 9 to 10 days after transfection, however, pMOV-ψ^--transfected cells produce infectious virus. Thus pMOV-ψ^- has a defect that can be repaired in transfected NIH/3T3 cells, presumably by recombination with a sequence normally present in the cells. Cell lines with pMOV-ψ^- stably integrated into chromosomal DNA produce reverse-transcriptase-containing particles that lack detectable M-MuLV RNA but the cells efficiently complement replication-defective, packagable retroviruses. Thus pMOV-ψ^- has a defect in the packaging of genomic RNA into virions but can provide in *trans* the products necessary for virion production. The deletion in pMOV-ψ^- appears to define a site required in *cis* for packaging of MuLV RNA into virions. Cell lines carrying pMOV-ψ^- can be used to produce helper-free stocks of natural or synthetic defective retroviruses.

Introduction

Retrovirus RNA is packaged into virions as a 70S dimer of two identical, capped and polyadenylated 35S RNAs (for review see Varmus, 1982). How the virus achieves selective packaging of these RNAs and excludes other mRNAs remains one of the more obscure aspects of the retrovirus life cycle. One possible explanation is that there is a site on the viral genome which interacts with a virion protein to direct specifically the packaging of the RNA. If such a site were deleted from a viral genome, the result would be a *cis* defect preventing the encapsidation of genomic RNA. Such a mutant, however, should still be capable of directing the synthesis of all viral proteins.

In searching for a *cis* packaging signal we noted that retroviruses package full-length RNA but not the spliced al., 1978; Levin and Seidman, 1979; Stoltzfus and Kuhnert, 1979). The ability of the virus to differentiate between these RNAs suggested that an essential part of a packaging signal might be in the region spliced out to form the *env* mRNA. If such a site was outside the coding region for Pr65^{gag} and reverse transcriptase, it would have to be 3′ of the presumed splice donor site and 5′ of the AUG that initiates translation of Pr65^{gag}.

A second clue to the location of a packaging site came from the study of a Rous sarcoma virus-transformed quail cell line that is deficient in packaging viral RNA (Linial et al., 1978). This mutant has a 150 bp deletion somewhere between 300 and 600 bp from the left end of the provirus thereby implicating once again the region upstream from the start of Pr65^{gag} as important for packaging (Shank and Linial, 1980). A complicating property of this mutant is that it is *trans*-dominant: wild-type genomes are also inefficiently packaged upon superinfection, suggesting that a second mutation exists in some "packaging factor" which relaxes the normal specificity for Rous sarcoma virus RNA (Linial et al., 1978; Linial, 1981).

In this study, we describe the deletion from Moloney murine leukemia virus (M-MuLV) of approximately 350 bp between the left long terminal repeat (LTR) and the start codon for Pr65^{gag}. The resulting mutant has a *cis*-active deficiency for packaging of genomic RNA and can be used to generate pure stocks of defective retroviruses, including recombinant retroviral vectors. Watanabe and Temin (1982) have provided evidence for a similar site in an avian retrovirus.

Results

Construction of pMOV-ψ^-

To examine the function of sequences upstream from the start of Pr65^{gag} and downstream from the *env* mRNA splice donor site in M-MuLV, we deleted these sequences from a cloned DNA representation of the genome. As detailed in Figure 1, the deletion was made from a Bal I site to a Pst I site, and a Hind III site was generated at the point of deletion. Based on the published sequence of M-MuLV (Shinnick et al., 1981), the size of the deletion was 351 bp. The 5′ deletion endpoint was only 6 bp from the presumed donor site for the *env* mRNA splice, and the 3′ deletion endpoint was approximately 50 bp from the start codon for Pr65^{gag} (Figure 2). The resulting deleted plasmid is termed pMOV-ψ^-.

Virus Spread Is Defective after Transfection of pMOV-ψ^-

To determine whether pMOV-ψ^- generated transmissible virus, it and its wild-type counterpart, pMOV-ψ^+, were independently transfected into NIH/3T3 cells, and reverse transcriptase in the culture supernatants was assayed over time. By 4 days after transfection, the medium from a culture transfected with pMOV-ψ^+ had 3-fold more reverse

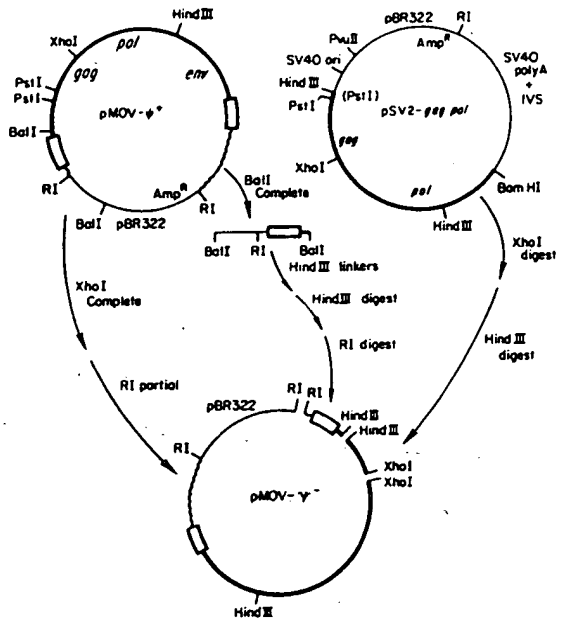


Figure 1. Construction of pMOV-ψ^-

pBR322 and SV40 sequences are represented by thin lines, MuLV sequences are represented by thick lines, and mouse sequences flanking the provirus are represented by wavy lines. The long terminal repeats of MuLV are indicated by open boxes. IVS: intervening sequence of SV40 small tumor antigen.

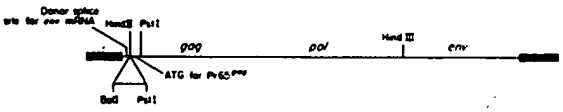


Figure 2. Schematic Representation of the Provirus Present in pMOV-ψ^-

The long terminal repeats are represented as closed boxes, and internal MuLV sequences are represented as a thin line. The 350 bp present between the Bal I and Pst I sites were deleted and a Hind III site was introduced at the site of deletion. This new Hind III site is six nucleotides from the apparent donor splice site for the *env* mRNA and 55 nucleotides from the ATG for $Pr65^{gag}$ (Shinnick et al., 1981).

crease in activity as a result of efficient spread of virus from the initially transfected cells to all cells on the plate. In contrast, for 4 to 6 days following transfection of pMOV-ψ^-, no detectable reverse transcriptase activity was observed in the medium, suggesting that virus spread was initially defective. Uninfected cells had less than 5% of the activity of MOV-1.

Within 9 days of transfection of pMOV-ψ^-, large amounts of reverse transcriptase activity developed in the culture supernatant (Figure 3), suggesting that virus spread eventually occurred. The reverse transcriptase activity present by day 9 could have been due either to the slow, attenuated spread of mutant virus or to the generation of a nondefective virus by recombination with a cellular sequence capable of providing the missing function. To differentiate between these possibilities, various dilutions of culture supernatants from cells transfected 9 days previously with pMOV-ψ^+ or pMOV-ψ^- were used to infect

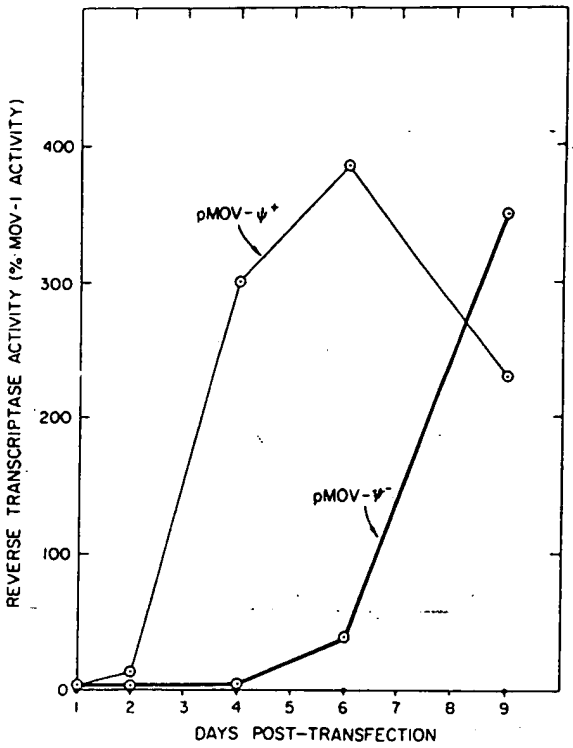


Figure 3. Reverse Transcriptase Activities of Culture Supernatants as a Function of Time After Transfection with Either pMOV-ψ^+ or pMOV-ψ^-

Solid dots represent background activity of untransfected NIH/3T3 cells. Thick line: pMOV-ψ^-. Thin line: pMOV-ψ^+. All reverse transcriptase activities are represented as a percentage of the activity produced by a standard MuLV producer cell line, MOV-1.

NIH/3T3 cells. High levels (3–4 times MOV-1 activity) of reverse transcriptase activity were observed 2 days after infection with only 10 μl of both supernatants (data not shown), demonstrating the presence of nondefective virus in the culture supernatants of cells transfected with pMOV-ψ^- 9 days previously.

pMOV-ψ^- Provides All Viral Functions Required in *Trans*

To determine whether the deletion made in pMOV-ψ^- affected functions required only in *cis* by the virus, helper activity was measured by using a recombinant retrovirus genome derived from Moloney sarcoma virus (MSV) capable of expressing the E. coli gene encoding xanthine-guanine phosphoribosyltransferase (XGPRT) (R. Cone, R. Mann, D. Baltimore, and R. C. Mulligan, unpublished data) (Figures 4A, 4B). This gene, designated *gpt*, is a dominant-acting selectable marker in mammalian cells (Mulligan and Berg, 1980). The plasmid (pMSVgpt) encodes no intact retroviral proteins yet retains the region deleted in pMOV-ψ^-. pMSVgpt also has the polyoma virus early region to increase its copy number in NIH/3T3 cells early after transfection. To assess the helper activity of pMOV-ψ^-, we cotransfected pMSVgpt and a 10-fold excess of either

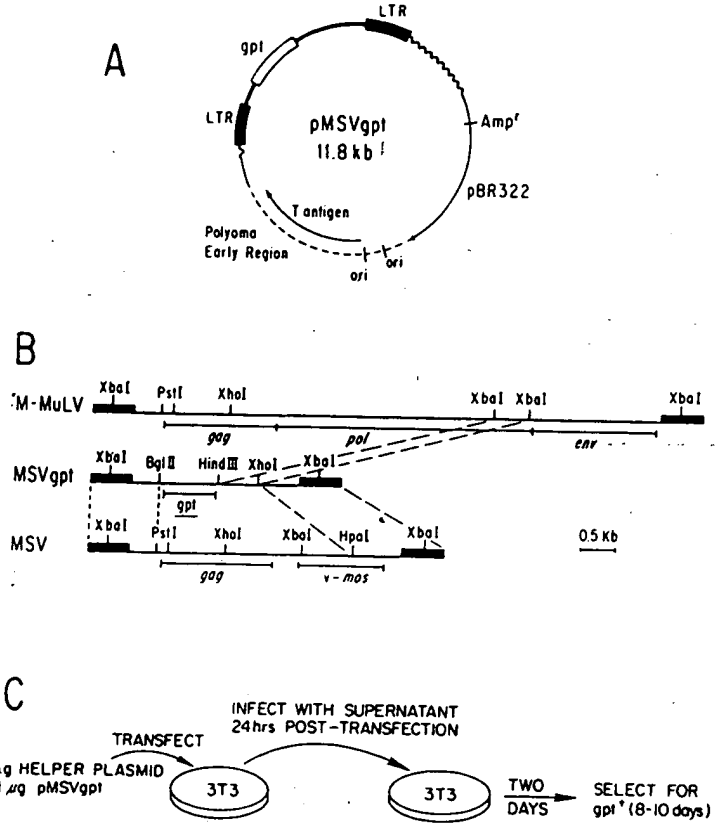


Figure 4. Structure of pMSVgpt and Its Use to Assay the Helper Activity of Retroviral Constructs

(A) Recombinant plasmid pMSVgpt is shown. Sequences derived from retroviruses are indicated by thick lines, with the long terminal repeats (LTRs) as filled boxes. The *gpt* gene (Mulligan and Berg, 1980) is indicated by an open box. Mink cellular DNA flanking the recombinant retrovirus is indicated by a wavy line. The polyoma virus early region (Magnusson and Nilsson, 1977) (base pairs 2932 to 4632) is indicated as a dashed line, and pBR322 sequences derived from the "poison minus" derivative (Lusky and Botchan, 1981) are indicated by a thin line.

(B) Structure and origin of the provirus segment of pMSVgpt showing that no intact retroviral genes are present. The reason for including the Xba I to Xba I fragment of M-MuLV in pMSVgpt is irrelevant to the present use of the plasmid. The MSV restriction map is from Vande Woude et al. (1979).

(C) Protocol for assessing the helper activity of retroviral constructs.

pMOV-ψ^+ or pMOV-ψ^- into NIH/3T3 cells (Figure 4C). After 24 hr, culture supernatants of the transfected cells were used to infect a second plate of NIH/3T3 cells. Two days following the infection, selection pressure for XGPRT was applied, and 8–10 days later *gpt*$^+$ colonies (colonies containing the *gpt* gene) were counted. pMOV-ψ^+ yielded 2.5 $\times$ 10^4 *gpt*$^+$ colony forming units (cfu) per milliliter of culture supernatant, and pMOV-ψ^- yielded 5 $\times$ 10^3 *gpt*$^+$ cfu per milliliter (Table 1). No *gpt*$^+$ cfu were detectable when pMSVgpt was transfected by itself. Thus pMOV-ψ^- could efficiently provide all functions in *trans* required for viable virus production.

Table 1. Assay of helper activity of pMOV-ψ^+ and pMOV-ψ^-

Helper Plasmid	*gpt*$^+$ (cfu/ml)*
pMOV-ψ^+	2.5 $\times$ 10^4
pMOV-ψ^-	5 $\times$ 10^3
none	<10

* *gpt*$^+$ colony forming units per milliliter of transfected cell culture supernatant 24 hr after transfection.

Construction of Cell Lines With pMOV-ψ^- Stably Integrated and Expressed

To explore further the properties of pMOV-ψ^-, we decided to establish cell lines containing the ψ^- mutant stably integrated into the chromosome. These lines were made by the cotransfection of pMOV-ψ^- and pSV2gpt, a SV40 hybrid vector capable of XGPRT expression (Mulligan and Berg, 1980). Cells from *gpt*$^+$ colonies obtained in this way were cloned and established into 3 lines, ψ-1, ψ-2 and ψ-3. All three lines had detectable reverse transcriptase activity in the culture supernatant (Table 2), indicating that they all contained sufficient MuLV information to express and bud reverse-transcriptase-containing particles. Only ml of culture supernatant from either ψ-1 or ψ-2 onto NIH/3T3 cells yielded no reverse transcriptase producer cells even after 3 weeks of culture. Thus the ψ-3 cells were producing a nondefective virus, whereas the ψ-1 and ψ-2 cells were not.

Viral RNA in ψ-2 Cells Is Poorly Packaged int Viri ns

Because the ψ^- mutant appeared defective in a function required only in *cis*, we suspected ψ^- RNA of missing sequences important for its packaging into virus particles. To address this question directly, total cellular RNA was purified from the MOV-1 (an M-MuLV producer) and ψ-2 cell lines, and viral RNA was extracted from purified particles released from the MOV-1 and ψ-2 cells. The MOV-1 and ψ-2 viral RNAs were derived from virion preparations

Table 2. Properties of cell lines stably transfected with pMOV-ψ^-

Helper Line	Reverse Transcriptase Activity in Supernatant[a]	Infectivity[b]	*gpt*$^+$ (cfu/ml) 24 hr after Transfection with pMSVgpt[c]
MOV-1	100.0	+	5×10^5
ψ-1	2.0	−	60
ψ-2	35.0	−	10^4
ψ-3	131.4	+	10^6
NIH/3T3	0.9	−	<10

[a] In arbitrary units with the wild-type MuLV producer line, MOV-1, set equal to 100.0.
[b] Infectivity was judged by infecting NIH/3T3 cells with 1 ml of the culture supernatant and measuring reverse transcriptase activity for at least two weeks following the infection.
[c] *gpt*$^+$ colony forming units per milliliter of culture supernatant of cells transfected with pMSVgpt.

activity and p30 protein by measuring their concentration in portions of the purified particles just prior to RNA extraction. Both cellular and viral RNAs from MOV-1 and ψ-2 were fractionated by electrophoresis through a 1% agarose–formaldehyde gel, transferred to a nitrocellulose filter, and probed with a nick-translated plasmid containing the entire M-MuLV genome (Figure 5A). Full-length genomic RNA could be detected in MOV-1 cells (8.2 kb, lane 1) and in ψ-2 cells (7.9 kb, lane 4), as could the subgenomic 2.9 kb *env* mRNA (lanes 1 and 4). The 8.2 kb genomic RNA could be readily detected in RNA preparations from MOV-1 virions (lanes 2 and 3), but no hybridizable RNA was detected in RNA preparations from an equivalent number of ψ-2 particles (lanes 5 and 6). In Figure 5A, lane 6 contains RNA from 10-fold more ψ-2 particles than were used for the MOV-1 particles in lane 3. Also, exposure of the autoradiogram shown in Figure 5A for a 10-fold longer time still revealed no hybridizable bands in lanes 5 and 6 (data not shown). Thus particles released from ψ-2 cells contain at most 1% of the viral RNA found in MOV-1 particles.

Further quantitative data were obtained by spotting 2-fold dilutions of MOV-1 and ψ-2 viral RNA preparations onto a nitrocellulose filter and probing it with a nick-translated plasmid containing the entire MuLV genome (Figure 5B). The amount of RNA spotted was again normalized to the number of particles present prior to RNA extraction by measuring reverse transcriptase activity and p30 concentration. From this analysis, it appeared that ψ^- viral RNA was packaged into virus particles with an efficiency much less than 1% that of wild-type viral RNA.

ψ-2 Can Efficiently Package Other Retroviral Genomes into Infectious Particles

We next examined whether the pMOV-ψ^--containing cell lines could package a highly defective retrovirus genome into virus particles and bud the particles into the culture supernatant without also budding helper virus. To this end, pMSVgpt was transfected into ψ-1, ψ-2, ψ-3, MOV-1 and

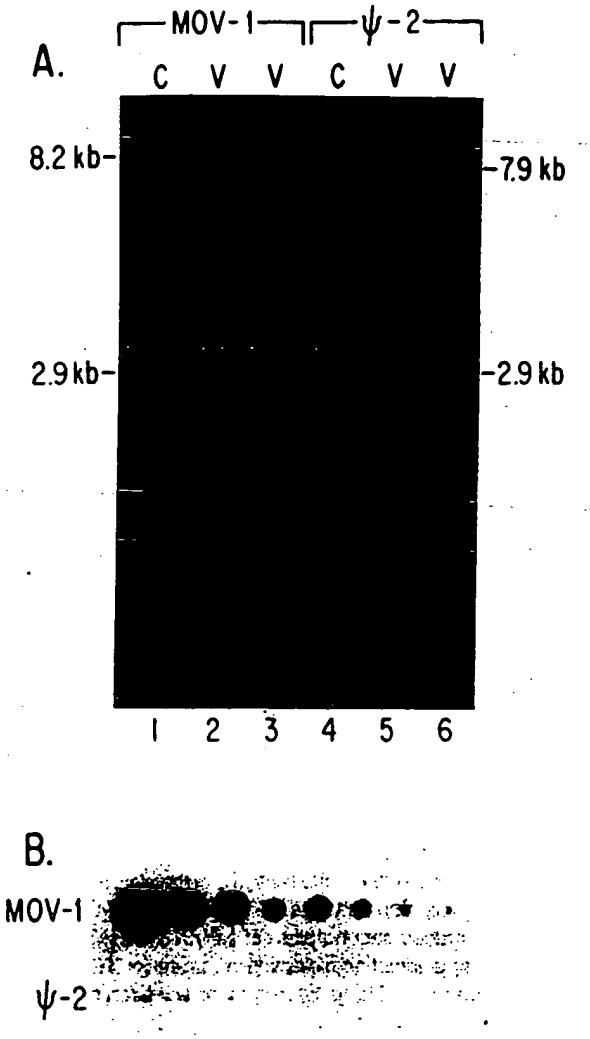


Figure 5. Hybridization Analysis of Cellular and Viral MuLV RNAs From MOV-1 and ψ-2 Cells

(A) Polyadenylated cellular RNA and viral RNA from MOV-1 and ψ-2 were fractionated by electrophoresis through a 1% agarose–formaldehyde gel, transferred to a nitrocellulose filter and probed with a nick-translated plasmid containing the entire MuLV genome. (Lane 1) 3 μg polyadenylated MOV-1 cellular RNA; (lane 2) viral RNA from purified MOV-1 particles; (lane 3) 10-fold the MOV-1 viral RNA present in lane 2; (lane 4) 3 μg polyadenylated ψ-2 cellular RNA; (lane 5) viral RNA from purified ψ-2 particles; (lane 6) 10-fold the ψ-2 viral RNA present in lane 5. Lanes 2 and 5 contain RNA extracted from equivalent numbers of MOV-1 and ψ-2 particles as determined by both reverse transcriptase activities and p30 concentrations.
(B) Sequential 2-fold dilutions of viral RNA obtained from purified MOV-1 virus or ψ-2 particles were spotted onto a nitrocellulose filter and probed with a nick-translated plasmid containing the entire MuLV genome. As in (A) the amount of viral RNA used was normalized to reverse transcriptase activities and p30 concentrations present in the particles prior to RNA extraction.

NIH/3T3 cells, and 24 hr later the culture supernatants were used to infect fresh NIH/3T3 cells. Two days after the infection, selection pr ssure for XGPRT was applied, and 8–10 days later *gpt*$^+$ colonies were counted. Cell line ψ-1 produced only 60 *gpt*$^+$ cfu/ml of culture supernatant

but ψ-2 yielded 10^4 cfu/ml (Table 2). Mov-1 and ψ-3 produced 5×10^5 and 10^5 cfu/ml, respectively. Furthermore, cells rendered *gpt*$^+$ by using supernatants from transfected ψ-1 or ψ-2 cells were producers of neither reverse transcriptase nor *gpt*$^+$ cfu and remained negative for both after 4 weeks of culture. In contrast, cells rendered *gpt*$^+$ using supernatants from transfected MOV-1 or ψ-3 cells were immediately producers of reverse transcriptase and *gpt*$^+$ cfu.

Discussion

By deleting 350 bp from a M-MuLV DNA clone, we have created a mutant, pMOV-ψ^-, that cannot itself be propagated as a virus, but that still efficiently provides all of the obligate *trans* retroviral functions. Upon transfection of cloned recombinant retroviruses, such as pMSVgpt, into cell lines expressing pMOV-ψ^-, such as ψ-2, high titers of infectious retroviral particles are produced which contain the recombinant genome. These particles are able to transmit the recombinant genome efficiently via the retrovirus life cycle. In contrast, the ψ^- genome is not detectably packaged into particles and therefore helper virus is not produced. Thus the defect in pMOV-ψ^- appears to be a *cis*-active defect and probably involves a packaging site on the RNA, presumably one recognized by a viral protein, rather than a region encoding a protein needed for packaging. Deletion of sequences in spleen necrosis virus at a similar position to that described here has been shown to result in a *cis*-active defect (Watanabe and Temin, 1982).

Murine retroviruses often encode a glycosylated form of their $Pr65^{gag}$ protein that is longer at the N-terminus than $Pr65^{gag}$ (Edwards and Fan, 1979, 1980). The pMOV-ψ^- clone contains a deletion of all of the information that might encode such an N-terminal extension and yet provides all of the needed helper virus functions. The 350 bp deletion also removes all three AUGs that are present between the 5′ LTR and the AUG used for $Pr65^{gag}$ initiation. It would thus appear that, at least in fibroblasts, the glycosylated *gag* precursor is not needed for efficient budding of infective progeny, although this has not been tested directly.

Dimerization and Packaging of Virion RNA

Electron microscopic studies of 70S virion RNA have suggested that a site approximately 300 nucleotides from the 5′ end of the 35S RNA molecule may be important for genome dimerization (Bender et al., 1978). This region is within the 350 bp deletion of pMOV-ψ^-. If dimerization is necessary for packaging, then the defect in pMOV-ψ^- may be a consequence of the deletion of the dimerization site. Studies by Cheung et al. (1972) and Canaani et al. (1973), however, suggested that the packaging and dimerization processes are at least temporally distinct, because they found that freshly budded virions of Rous sarcoma virus contained only the monomer RNA and that dimer formation followed budding. Smaller mutations will have to be constructed to examine whether packaging and dimerization

Generation of Nondefective Viruses

When pMOV-ψ^- DNA is transfected into mouse cells, nondefective virus appears after a lag of 4 to 6 days. The slow appearance of nondefective virus, and the ease of isolating cloned cell lines that harbor pMOV-ψ^- but do not generate nondefective virus, suggest that only a minority of transfected cells yield nondefective virus. We assume that recombination between some cellular sequence and pMOV-ψ^--derived DNA or RNA is responsible for the generation of nondefective virus, but we have yet to provide direct evidence for this. Likely candidates for providing the packaging site sequence are the family of abundant, packagable, virus-like 30S RNAs found in mouse cells called VL30 RNAs (Sherwin et al., 1978; Besmer et al., 1979; Scolnick et al., 1979) and the many retrovirus-like DNA elements found in normal mouse cells (reviewed in Coffin, 1982). We are currently analyzing the structure of the nondefective viruses as well as determining whether these genomes can be generated in other cell types.

Use of ψ-2

In a number of viral systems it has proved useful to design cell lines capable of providing certain viral functions. Such cell lines can then be used as hosts for growth of viruses defective for production of the proteins already present in the cells. For instance, the hr-t mutants of polyoma virus were isolated in this way (Benjamin, 1970), early mutants of SV40 can be grown in COS cells (Gluzman, 1981), and early mutants of adenovirus can be grown in 293 cells (Graham et al., 1977). The ψ-2 cell line described here is an analogous cell line for defective retroviruses. With such a line, stocks of natural or engineered defective retroviruses, such as pMSVgpt and others (Mulligan, 1983), can be obtained totally free of detectable helper virus. This tool enhances the utility of retroviral vectors (Wei et al., 1981; Tabin et al., 1982; Mulligan, 1983) because pure stocks of the engineered viruses can be used to introduce genes into cells via the retroviral life cycle without the recipient cell becoming a retrovirus producer. Furthermore, genes cloned into retroviral genomes could, in principle, be introduced into the germ line of mice without also introducing helper virus, in much the same way as has been demonstrated for MuLV (Jaenisch, 1976, 1977; Jaenisch et al., 1981), thus avoiding side effects inherent in having helper virus present.

Experimental Procedures

Cells and Viruses

The NIH/3T3 cell line was grown in Dulbecco's modified Eagle medium containing 10% calf serum. The MOV-1 cell line was derived by transfection of pMOV-ψ^+ (actually pMOV-9, a gift from R. Jaenisch) into NIH/3T3 cells followed by cloning. Cells to be cloned were trypsinized, counted, and seeded into 96-well cloning trays at 0.3 cells per well (in a volume of 0.2 ml per well) and allowed to grow for 10 to 12 days. Selection medium for *gpt*$^+$ was as described (Mulligan and Berg, 1981). Virus infections were performed in the presence of 8 μg/ml polybrene for 1–2 hr.

Reverse Transcriptase Assays

Reverse transcriptase assays were performed as described (Goff et al., 1981). Subconfluent cells were fed for 12 hr before being assayed to

DNA Transfection

Twenty-four hours after 10^6 cells were seeded on 100 mm plates, the cultures were transfected with 10 μg of plasmid DNA by using the procedure of Graham and Van der Eb (1973) as modified by Parker and Stark (1979). In cases where selection for transfected cells was applied, cells were split 1:20 into selective media 48 hr after transfection. Unselectable DNA was cotransfected in 10-fold molar excess over 1 μg of selectable DNA.

Preparation and Analysis of RNA

Total cellular RNA was prepared by the method of Chirgwin et al. (1979), and polyadenylated RNAs were selected with oligo(dT)-cellulose as described by Quintrell et al. (1980). Virion RNA was prepared by extraction from virus particles banded in sucrose step gradients (Shields et al., 1978). RNA was size-fractionated by electrophoresis through agarose–formaldehyde gels (Maniatis et al., 1982), transferred to nitrocellulose, and hybridized with ^{32}P-labeled DNA probes as described (Rigby et al., 1977). Quantitation of RNA by dot blots was done as described by Thomas (1980).

Construction of pMOV-ψ⁻

Three purified DNA fragments were ligated together to construct pMOV-ψ⁻ (Figure 1). The first was obtained by digesting pMOV-ψ⁺ (Chumakov et al., 1982) with Xho I to completion, followed by partial digestion with Eco RI. The fragment extending from the Xho I site at 2.0 U in MuLV, through the 3′ LTR, 3′ mouse flanking sequence, all of pBR322, and ending at the Eco RI site was purified from an agarose gel after electrophoretic separation (Vogelstein and Gillespie, 1979). The second fragment was obtained by digestion of pMOV-ψ⁺ with Bal I to completion followed by purification of the fragment extending from the Bal I site in pBR322 through 5′ mouse flanking sequence and 5′ LTR to the Bal I site located at 0.7 U of MuLV. Hind □ linkers (Collaborative Research) were then blunt-ligated to this fragment with T4 DNA ligase, and the fragment was digested with excess Hind □ and Eco RI. The LTR-containing fragment was purified from an agarose gel after electrophoretic separation. The third fragment present in the final ligation reaction was obtained from pSV2-*gag/pol* (R. Mulligan, unpublished plasmid) where the *gag/pol* region of MuLV had been subcloned into pSV2 (Mulligan and Berg, 1980). pSV2-*gag/pol* was digested to completion with Xho I and Hind □ and the fragment extending from the Hind □ site (changed from the Pst I site at 1.0 U of MuLV) to the Xho I site at 2.0 U of MuLV was purified from an agarose gel following electrophoretic separation. These three DNA fragments were then mixed in equimolar amounts at a total DNA concentration of 50 μg/ml in ligase buffer (50 mM Tris-HCl [pH 7.8], 10 mM $MgCl_2$, 20 mM dithiothreitol, 1.0 mM ATP, 50 μg/ml bovine serum albumin) and incubated with T4 DNA ligase for 18 hr at 15°C. *E. coli* HB101 was transfected with the ligated DNA, and ampicillin-resistant transfectants were obtained. The plasmid DNA obtained from a number of transformants was screened for the desired structure by digestion with appropriate restriction endonucleases and electrophoresis through agarose gels (Davis et al., 1980).

Acknowledgments

R. C. M. wishes to acknowledge the support of Dr. J. T. Potts, Jr., during the early part of these studies.

This work was supported by grants from the National Cancer Institute and the National Institutes of Health Program Project Grant (CA 26712-03). R. C. M. is a McArthur Foundation Prize Fellow. D. B. is an American Cancer Society Research Professor.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 4, 1983; revised March 7, 1983

References

Bender, W., Chien, Y.-H., Chattopadhyay, S., Vogt, P. K., Gardner, M. B., and Davidson, N. (1978). High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. J. Virol. *25*, 888–896.

Benjamin, T. L. (1970). Host range mutants of polyoma virus. Proc. Nat. Acad. Sci. USA *67*, 394–399.

Besmer, P., Olshevsky, U., Baltimore, D., Dolberg, D., and Fan, H. (1979). Virus-like 30S RNA in mouse cells. J. Virol. *29*, 1168–1176.

Canaani, E., Helm, K. V. D., and Duesberg, P. (1973). Evidence for 30-40S RNA as precursor of the 60–70S RNA of Rous sarcoma virus. Proc. Nat. Acad. Sci. USA *70*, 401–405.

Cheung, K.-S., Smith, R. E., Stone, M. P., and Jokik, W. K. (1972). Comparison of immature (rapid harvest) and mature Rous sarcoma virus particles. Virology *50*, 851–864.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18*, 5294.

Chumakov, I., Stuhlmann, H., Harbers, K., and Jaenisch, R. (1982). Cloning of two genetically transmitted Moloney leukemia proviral genomes: correlation between biological activity of cloned DNA and viral genome activation in the animal. J. Virol. *42*, 1088–1098.

Coffin, J. (1982). Endogenous Viruses. In RNA Tumor Viruses, Weiss, R., Teich, N., Varmus, H., and Coffin, J., eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 1147–1155.

Davis, R. W., Thomas, M., Cameron, J., St. John, T. P., Scherer, S., and Padgett, R. (1980). Rapid DNA isolations for enzymatic and hybridization analysis. Meth. Enzymol. *65*, 404–411.

Edwards, S. A., and Fan, H. (1979). *gag*-related polyproteins of Moloney murine leukemia virus: evidence for independent synthesis of glycosylated and unglycosylated forms. J. Virol. *30*, 551–563.

Edwards, S. A., and Fan, H. (1980). Sequence relationship of glycosylated and unglycosylated *gag* polyproteins of Moloney murine leukemia virus. J. Virol. *35*, 41–51.

Gerwin, B. I., and Levin, J. G. (1977). Interactions of Murine leukemia virus core components: characterization of reverse transcriptase packaged in the absence of 70S genomic RNA. J. Virol. *24*, 478–488.

Gluzman, Y. (1981). SV40-transformed simian cells support the replication of early SV40 mutants. Cell *23*, 175–182.

Goff, S., Traktman, P., and Baltimore, D. (1981). Isolation and properties of Moloney murine leukemia virus mutants: use of a rapid assay for release of virion reverse transcriptase. J. Virol. *38*, 239–248.

Graham, F. L., Smiley, J., Russell, W. C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. *36*, 59–72.

Graham, R., and Van der Eb, A. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virol. *52*, 456–467.

Jaenisch, R. (1976). Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc. Nat. Acad. Sci. USA *73*, 1260–1264.

Jaenisch, R. (1977). Germ line integration of Moloney leukemia virus: effect of homozygosity at the M-MuLV locus. Cell *12*, 691–696.

Jaenisch, R., Jahner, D., Nobis, P., Simon, I., Lohler, J., Harbers, K., and Grotkopp, D. (1981). Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell *24*, 519–529.

Levin, J., and Seidman, J. (1979). Selective packaging of host tRNA's by murine leukemia virus particles does not require genomic RNA. J. Virol. *29*, 328–335.

Linial, M. (1981). Transfer of defective avian tumor virus genomes by a Rous sarcoma virus RNA packaging mutant. J. Virol. *38*, 380–382.

Linial, M., Medeiros, E., and Hayward, W. S. (1978). An avian oncovirus mutant (SE 21Q1b) deficient in genomic RNA: biological and biochemical characterization. Cell *15*, 1371–1381.

Lusky, M., and Botchan, M. (1981). Inhibition of SV40 replication in simian cells by pBR322 sequences. Nature *293*, 79–81.

Magnusson, G., and Nilson, M.-G. (1977). Multiple free viral DNA copies in polyoma virus-transformed mouse cells surviving productive infection. J. Virol. *22*, 646–653.

Maniatis, T., Fritsch, P., and Sambrook J. (1982). Molecular Cloning, A

Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 202–203.

Mulligan, R. C. (1983) Construction of highly transmissible mammalian cloning vehicles derived from murine retroviruses. In Experimental Manipulation of Gene Expression, M. Inouye, ed. (New York: Academic Press), in press.

Mulligan, R. C., and Berg, P. (1980). Expression of a bacterial gene in mammalian cells. Science *209*, 1422–1427.

Mulligan, R. C., and Berg, P. (1981). Selection for animal cells that express the *E. coli* gene coding for xanthine–guanine phosphoribosyltransferase. Proc. Nat. Acad. Sci. USA *78*, 2072–2076.

Quintrell, N., Hughes, S. H., Varmus, H. E., and Bishop, J. M. (1980). Structure of viral DNA and RNA in mammalian cells infected with avian sarcoma virus. J. Mol. Biol. *143*, 363–393.

Parker, B. A., and Stark, G. R. (1979). Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. *31*, 360–369.

Rigby, P. W. J., Dickmann, M., Rhodes, C., and Berg, P. (1977). Labeling in vitro by nick translation with DNA polymerase I. J. Mol. Biol. *113*, 237–251.

Rothenberg, E., Donoghue, D., and Baltimore, D. (1978). Analysis of a 5′ leader sequence on murine leukemia virus 21S RNA: heteroduplex mapping with long reverse transcriptase products. Cell *13*, 435–451.

Scolnick, E. M., Vass, W. C., Howk, R. S., and Duesberg, P. H. (1979). Defective retrovirus-like 30S RNA species of rat and mouse cells are infectious if packaged by type C helper virus. J. Virol. *29*, 964–972.

Shank, P., and Linial, M. (1980). Avian oncovirus mutant (SE21Q1b) deficient in genomic RNA: characterization of a deletion in the provirus. J. Virol. *36*, 450–456.

Sherwin, S. A., Rapp, U. R., Benveniste, R. E., Sen, A., and Todaro, G. J. (1978). Rescue of endogenous 30S retroviral sequences from mouse cells by baboon type C virus. J. Virol. *26*, 257–264.

Shields, A., Witte, O. N., Rothenberg, R., and Baltimore, D. (1978). High frequency of aberrant expression of Moloney murine leukemia virus in clonal infections. Cell *14*, 601–609.

Shinnick, T. M., Lerner, R. A., and Sutcliffe, J. G. (1981). Nucleotide sequence of Moloney murine leukemia virus. Nature. *293*, 543–548.

Stoltzfus, C. M., and Kuhnert, L. K. (1979). Evidence for the identity of shared 5′-terminal sequences between genome RNA and subgenomic mRNA's of B77 avian sarcoma virus. J. Virol. *32*, 536–545.

Tabin, C., Hoffman, J., Goff, S., and Weinberg, R. (1982). Adaptation of a retrovirus as a eucaryotic vector transmitting the Herpes simplex virus thymidine kinase gene. Mol. Cell. Biol. *2*, 426–436.

Thomas, P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Nat. Acad. Sci. USA *77*, 5201–5205.

Vande Woude, G. F., Osakarsson, M., Enquist, L. W., Nomura, S., Sullivan, M., and Fishinger, P. J. (1979). Cloning of integrated Moloney sarcoma proviral DNA sequences in bacteriophage λ. Proc. Nat. Acad. Sci. USA *76*, 4464–4468.

Varmus, H. E. (1982). Form and function of retroviral proviruses. Science. *216*, 812–820.

Vogelstein, B., and Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. Proc. Nat. Acad. Sci. USA *76*, 615–619.

Watanabe, S., and Temin, H. (1982). Encapsidation sequences for spleen necrosis virus, and avian retrovirus, aare between the 5′ long terminal repeat and the start of the *gag* gene. Proc. Nat. Acad. Sci. USA *79*, 5986–5990.

Wei, C.-M., Gibson, M., Spear, P. G., and Scolnick, E. M. (1981). Construction and isolation of a transmissible retrovirus containing the src gene of Harvey murine sarcoma virus and the thymidine kinase gene of Herpes simplex virus type 1. J. Virol. *39*, 935–944.

ITP-3

The EMBO Journal vol.14 no.17 pp.4276–4287, 1995

Septe

Activating point mutations in th common β subunit of the human GM-CSF, IL-3 and IL-5 r c ptors sugg st the involvement of β subunit dimerization and cell type-specific molecules in signalling

Brendan J.Jenkins, Richard D'Andrea and Thomas J.Gonda[1]

Hanson Centre for Cancer Research and Division of Human Immunology, Institute of Medical and Veterinary Science, Frome Road, Adelaide, SA 5000, Australia

[1]Corresponding author

We have combined retroviral expression cloning with random mutagenesis to identify two activating point mutations in the common signal-transducing subunit (hβc) of the receptors for human granulocyte–macrophage colony-stimulating factor (GM-CSF), interleukin (IL)-3 and IL-5 by virtue of their ability to confer factor independence on the haemopoietic cell line, FDC-P1. One mutation (V449E) is located within the transmembrane domain and, by analogy with a similar mutation in the *neu* oncogene, may act by inducing dimerization of hβc. The other mutation (I374N) lies in the extracellular, membrane-proximal portion of hβc. Neither of these mutants, nor a previously described mutant of hβc (FIΔ, which has a small duplication in the ε tracellular region), was capable of inducing factor independence in CTLL-2 cells, while only V449E could induce factor independence in BAF-B03 cells. These results imply that the extracellular and transmembrane mutations act by different mechanisms. Furthermore, they imply that the mutants, and hence also wild-type hβc, interact with cell type-specific signalling molecules. Models are presented which illustrate how these mutations may act and predict some of the characteristics of the putative receptor-associated signalling molecules.

Keywords: cytokine receptor superfamily/human GM-CSF receptor common β chain/oncogenic activation/polymerase chain reaction mutagenesis/retroviral expression cloning

Introduction

Granulocyte–macrophage colony-stimulating factor (GM-CSF) and interleukin (IL)-3 stimulate the proliferation, differentiation and functional activity of a wide variety of haemopoietic cells, including neutrophils, eosinophils, monocytes and early progenitor cells (reviewed by Metcalf, 1986; Clark and Kamen, 1987). In addition, the capacity f GM-CSF and IL-3 to stimulate the proliferation of eosinophil progenitors is also shared by IL-5 (reviewed by Sanderson, 1992). This functional overlap, as well as the cross-competition f r binding to receptors n the surface of human haem poietic cells (Lopez *et al.*, 1991), has a clear correlate in the structure and composition of the receptors for these three CSFs. The high-affinity receptors for human GM-CSF (hGMR), IL-3 (hIL-3R) and IL-5 (hIL-5R) are composed of ligand-specific α subunits (hGMRα, hIL-3Rα and hIL-5Rα) associated with a common β subunit (hβc). The α subunits bind their cognate factors with low affinity, whereas hβc alone does not bind any of these factors detectably, but is required, in association with the α subunits, to confer high-affinity binding (Gearing *et al.*, 1989; Hayashida *et al.*, 1990; Kitamura *et al.*, 1991a; Tavernier *et al.*, 1991). Moreover, hβc is essential for signal transduction (Kitamura *et al.*, 1991b; Kitamura and Miyajima, 1992; Sakamaki *et al.*, 1992) and, as a shared signal-transducing component, provides a molecular explanation for the overlapping biological characteristics of GM-CSF, IL-3 and IL-5.

Both subunits of the receptors for GM-CSF, IL-3 and IL-5 are members of a large family, often termed the cytokine receptor family (reviewed by Miyajima *et al.*, 1992), that includes the receptors for many haemopoietic growth factors. Receptors of this class are characterized by a conserved structure [including the hallmark WSXWS (Trp-Ser-Xaa-Trp-Ser) motif] in their extracellular domains, which was first predicted by Bazan (1990), and then refined by determination of the structure of the complex between human growth hormone (hGH) and its receptor (hGHR) (de Vos *et al.*, 1992). Unlike the other major class of growth factor receptors, the receptor tyrosine kinases, members of the cytokine receptor family do not appear to possess any intrinsic enzymatic activities that could account for intracellular signalling.

Despite the fact that β subunits and their associated α subunits lack an intrinsic tyrosine kinase activity, it has been thoroughly documented that tyrosine phosphorylation of a number of proteins, including the β subunits themselves, is rapidly induced following ligand binding (Duronio *et al.*, 1992; Sakamaki *et al.*, 1992; Hanazono *et al.*, 1993). At least two distinct regions of the intracellular domain of hβc are involved in the generation of separate intracellular responses (Sakamaki *et al.*, 1992; Sato *et al.*, 1993), which implies that hβc is associated with and utilizes multiple effector molecules. Indeed, phosphorylation is believed to be carried out by one or more receptor-associated tyrosine kinases such as JAK-2 (Silvennoinen *et al.*, 1993), Fes (Hanazono *et al.*, 1993) and Lyn (Torigoe *et al.*, 1992). However, the mechanisms by which the effector molecules are activated are not clear. While it is accepted that signalling is mediated by a complex comprising the ligand and both the α and β subunits, neither the precise stoichiometry nor the role of each subunit in signalling is clear. On one hand, the cytoplasmic portion of the α subunit is required for signalling by the GM-CSF and IL-5 receptors (Sakamaki *et al.*, 1992; Takaki *et al.*, 1993; Polotskaya *et al.*, 1994). On the other hand, studies with chimeric receptors suggest that dimerization f the intracellular portion of the β

Oxf rd University Press

subunit may be sufficient t initiate signal transduction (e.g. Sakamaki *et al.*, 1993; Takaki *et al.*, 1994; see Discussion).

Regardless of the precise mechanism by which they function, the GM-CSF, IL-3 and IL-5 receptors transduce proliferative signals, and it would theref re seem likely that their common signalling subunit, hβc, represents a potential target for oncogenic activation. Indeed, the oncogenic potential of several members of the cytokine receptor superfamily has already been demonstrated: the v-*mpl* oncogene of the myeloproliferative leukaemia virus (MPLV) encodes a truncated form of the recently characterized c-MPL receptor for thrombopoietin (Souyri *et al.*, 1990; Vigon *et al.*, 1992; Lok *et al.*, 1994), and a single amino acid substitution in the extracellular domain of the erythropoietin receptor confers factor-independent growth and tumorigenicity on factor-dependent haemopoietic cells (Yoshimura *et al.*, 1990). More recently, a constitutively active mutant of hβc (FIΔ), that arose spontaneously following retroviral transduction, was isolated by virtue of its ability to confer factor independence and tumorigenicity on a murine haemopoietic cell line (D'Andrea *et al.*, 1994).

Here we report the identification of two activating point mutations in hβc that were generated using a polymerase chain reaction (PCR)-based random mutagenesis procedure. By combining this procedure with a retroviral expression cloning system recently developed in our laboratory (Rayner and Gonda, 1994), these mutants were isolated by virtue of their ability to confer factor independence on a factor-dependent haemopoietic cell line. The ability of point mutations to activate hβc supports the possibility that such mutations could contribute to human leukaemia. In addition, the nature and properties of these mutants also have important implications for the mechanism of signalling by the receptors that utilize the hβc subunit.

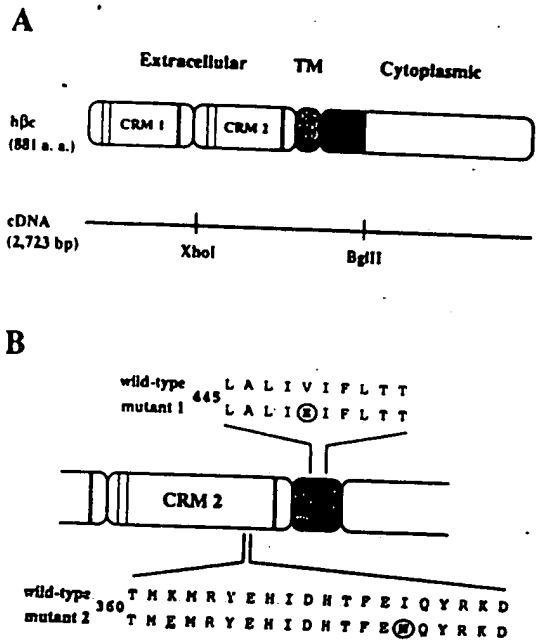


Fig. 1. Location of activating point mutations in the region of the β subunit of the human GM-CSF/IL-3/IL-5 receptors subjected to random mutagenesis. (A) Schematic illustration of hβc showing the two cytokine receptor modules (CRMs: Goodall *et al.*, 1993), the conserved cysteine residues (thin vertical lines) and the characteristic WSXWS motifs (thick vertical lines) [see Bazan (1990) for description of these elements]. Also shown are the transmembrane domain (hatched) and the minimal region essential for proliferative signalling in the cytoplasmic domain (Sakamaki *et al.*, 1992) (black shading). The positions in the hβc cDNA of the *XhoI* and *BglII* sites that delimit the fragment used for mutagenesis are shown underneath. (B) Location of point mutations in activated forms of hβc. The structural features of hβc are illustrated as in (A). The sequences of the regions containing mutations in mutants 1 and 2 (see text) are expanded; the activating mutations are circled and in bold type and the additional mutation in mutant 2 is underlined.

Results

Generation of a retroviral expression library of point-mutated hβc cDNA constructs

To investigate the oncogenic potential of hβc, we introduced point mutations randomly into a 940 bp segment of the hβc cDNA encoding the 191 membrane-proximal residues of the extracellular domain, the transmembrane domain and the first 92 residues of the cytoplasmic domain (Figure 1A). We identified this segment as a potential target for activating mutations because it encompasses several sequences that have been implicated in signalling and/or activation of cytokine receptors: (i) the membrane-proximal cytoplasmic region essential for signal transduction (D'Andrea *et al.*, 1991; Sakamaki *et al.*, 1992; Sato *et al.*, 1993; Ziegler *et al.*, 1993); (ii) the highly conserved WSXWS sequence that is a hallmark of the cytokine receptor superfamily (Bazan, 1990); (iii) the equivalent extracellular region of c-Mpl remaining in the v-Mpl oncoprotein (Souyri *et al.*, 1990; Vigon *et al.*, 1992); and (iv) the extracellular sequence duplicated in the constitutively active hβc mutant described by D'Andrea *et al.* (1994).

Our approach to identifying activating mutations in hβc was t c nstruct a bank of random point mutants and then insert these en masse into a retroviral expression vector. The resultant retroviruses were then used to infect a murine factor-dependent haemopoietic cell line (FDC-P1), following which these cells were selected for the ability to grow in the absence of factor. This procedure, which is outlined in Figure 2, is essentially a combination of the PCR mutagenesis technique described by Cadwell and Joyce (1992) and methodology previously developed in this laboratory (Rayner and Gonda, 1994).

As described in Materials and methods, we were able to define PCR conditions, based on those described by Cadwell and Joyce (1992), that resulted in the unbiased, random generation of mutations at the desired rate of ~0.2% (1 in 500 bp). The mutated hβc fragments were inserted directionally into pRUFNeo-hβc (see Materials and methods and Figure 2) from which the segment subjected to PCR mutagenesis had been excised. This resulted in the generation of a library, comprising 1.7×10^5 plasmid clones, of hβc cDNAs bearing point mutations in the targeted segment only ($h\beta c^{mut}$). Using procedures described previously (Rayner and Gonda, 1994), the plasmid DNA was used to generate a retroviral library estimated to contain 2.1×10^4 independent viral producer cl nes. Assuming that there was no verwhelming bias in the procedure, this should adequately represent all of the possible point mutations in the 940 bp hβc fragment.

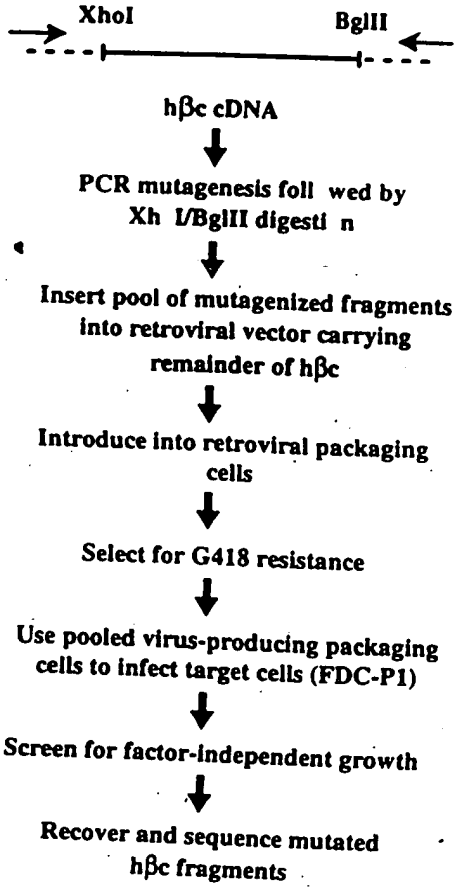


Fig. 2. Outline of the strategy used for generation and expression of hβc mutants. The arrows at the top represent the primers used for PCR amplification/mutagenesis of the segment of hβc shown in Figure 1A; these lie just outside the *BglII* and *XhoI* sites (see Materials and methods). Mutagenized fragments were ligated into the pRUFNeo-hβc plasmid which had been digested with *BglII* and *XhoI*. The resultant pool of plasmids was introduced into Ψ2 retroviral packaging cells as described by Rayner and Gonda (1994). Following G418 selection, the pooled virus-producing Ψ2 cells were used to infect FDC-P1 cells, which were subsequently selected for factor independence. The mutated hβc fragments were recovered from factor-independent FDC-P1 cells by PCR using the same primers as for the original mutagenesis/amplification.

Isolation of factor-independent FDC-P1 cells carrying constitutively active hβc mutants

FDC-P1 cells are dependent on mouse GM-CSF (mGM-CSF) or IL-3 (mIL-3) for growth, and normally die within days when starved of either factor. However, they do proliferate in response to low concentrations of human GM-CSF (hGM-CSF) or IL-3 (hIL-3) if the appropriate α subunit (hGMRα or hIL-3Rα) is co-expressed with hβc (our unpublished results). We therefore reasoned that infection of FDC-P1 cells with a retrovirus encoding a constitutively active form of hβc should induce factor-independent proliferation; this approach has been validated by the recent isolation of a constitutively active hβc mutant, FIΔ, expressed in factor-independent FDC-P1 cells (D'Andrea *et al.*, 1994).

Thus, to isolate factor-independent cells expressing constitutively active hβc point mutants present in the hβc expression library, FDC-P1 cells were infected by co-cultivation with the pool of $h\beta c^{mut}$ virus-producing Ψ2 cells (see Materials and methods). Parallel co-cultivati ns were also performed with uninfected Ψ2 cells and Ψ2 cells producing wild-type hβc retrovirus. After 1 week in liquid culture in the absence of factor, all four FDC-P1 cell pools infected with the mutant hβc retroviral library contained viable, proliferating cells, while no such cells were present in the control cultures. Factor independence was not a result of autocrine growth factor production as conditioned medium from the factor-independent cell pools did not support the growth of uninfected FDC-P1 cells (data not shown). A total of 31 factor-independent clones was isolated from the factor-independent liquid culture pools by agar plating; each of these was analysed further. Based on infection frequency and the number of independent mutants (detected by sequencing and/or restriction analysis; see also below and Figure 7), we estimate the frequency of factor-independent mutations to be ≥ 1 in 10^5.

Identification of activating mutations in constitutively active hβc mutants

During the course of recovering the mutated region of hβc from the factor-independent clones by PCR, it was discovered that 22/31 clones contained an additional *BglII* restriction site (as shown in Figure 7). Sequencing of one such clone (mutant 1) revealed that this was due to a T to A mutation at nucleotide 1374, which results in a change of the amino acid valine to glutamic acid at residue 449 (V449E) within the transmembrane domain (see Figure 1B). This is strikingly reminiscent of the activating point mutation (V664E) found by Bargmann *et al.* (1986) in the rat *neu* oncogene (see Discussion). Sequence analysis of one other clone that lacked the extra *BglII* site (mutant 2) revealed two point mutations: an A to G substitution at nucleotide 1112 which resulted in the substitution of lysine 362 to glutamic acid (K362E), and a T to A substitution at nucleotide 1149 which resulted in the substitution of isoleucine 374 to asparagine (I374N), both of which lie in the extracellular portion of the targeted region (see Figure 1B).

To confirm that the V449E mutation was indeed capable of activating hβc, and to identify which of the two mutations in mutant 2 was responsible for activation, all three mutations were re-created independently by site-directed mutagenesis. We also constructed two additional mutants, V449Q and V449D, based on other mutations known to activate c-*neu* (Bargmann and Weinberg, 1988; see also below). Following insertion into the pRUFNeo retroviral vector and transfection into Ψ2 packaging cells, these mutants, as well as wild-type hβc, were again introduced into FDC-P1 cells, which were then selected either for G418 resistance or for growth in medium without mGM-CSF. All of the viruses efficiently generated G418-resistant cells, which were subsequently analysed for hβc expression by antibody staining and flow cytometry. In these and subsequent experiments, we also included the previously described activated hβc mutant FIΔ, which contains a 37 amino acid duplication in the membrane-proximal portion of the extracellular domain (D'Andrea *et al.*, 1994). Figure 3 shows that a substantial proportion of FDC-P1 cells infected with each of these viruses, except that carrying the V449D mutant, expressed hβc.

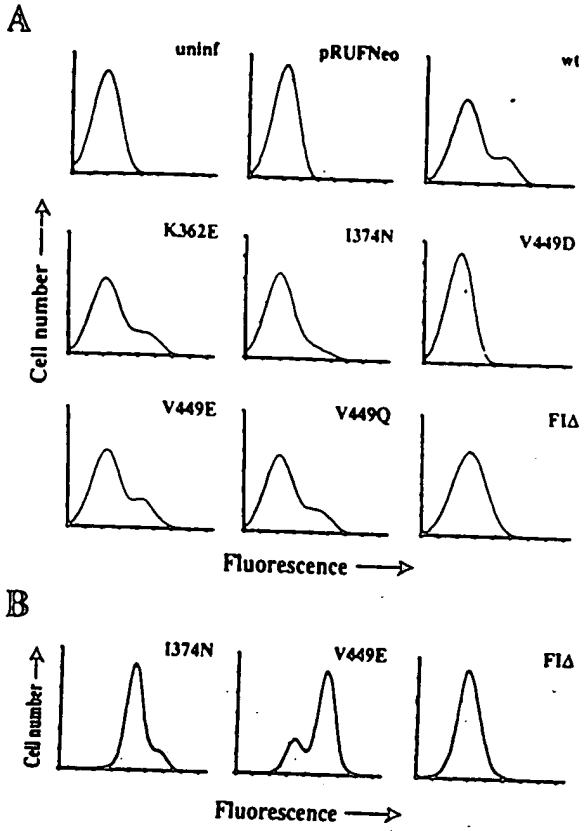


Fig. 3. Flow cytometric analysis of wild-type (wt) and mutant hβc expression on infected FDC-P1 cells. Dotted lines (.....) represent cells stained with an irrelevant control antibody and solid lines (—) indicate staining with anti-hβc antibody. Cell number and fluorescence are in arbitrary units; the latter is plotted on a logarithmic scale. (A) Cells selected for resistance to G418; also shown are analyses of uninfected cells and cells infected with the pRUFNeo vector alone. (B) Cells selected for growth in the absence of factor.

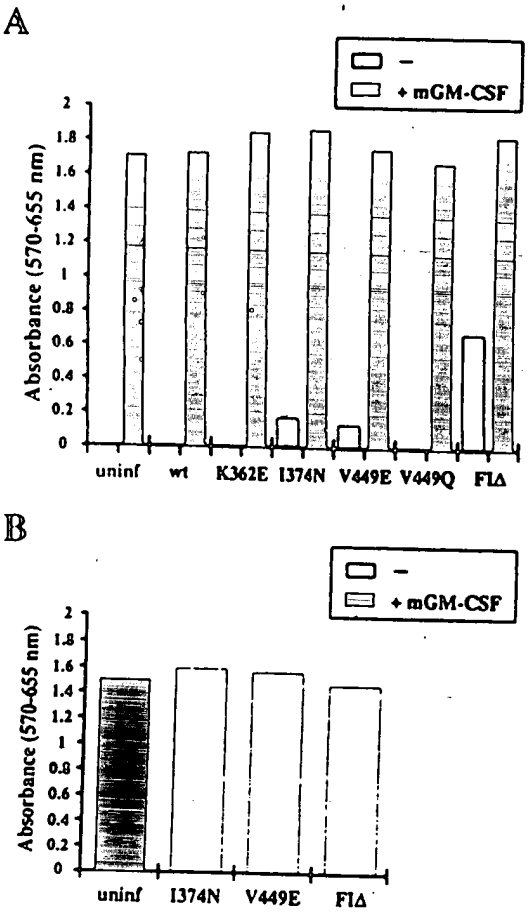


Fig. 4. Proliferation of FDC-P1 cells infected with hβc mutants in the presence and absence of murine GM-CSF. (A) Proliferation of cells selected and maintained prior to assay in mGM-CSF plus (except for the uninfected control cells) G418. Proliferation assays were carried out, as described in Materials and methods, in the presence or absence of mGM-CSF for uninfected FDC-P1 cells and cells infected with wild-type hβc (wt) or the indicated hβc mutants. (B) Proliferation of FDC-P1 cells infected with activated hβc mutants that were selected prior to assay by growth in factor-free medium. Proliferation of uninfected FDC-P1 cells in the presence of mGM-CSF is shown for comparison.

(Repeated attempts to obtain surface expression of V449D were unsuccessful.) However, only the I374N, V449E and, as expected, FIΔ mutants conferred factor independence on the FDC-P1 cells; the K362E, V449Q or (not surprisingly) the V449D mutants all failed to do so (Figure 4A; the data for the non-expressing V449D mutant is not shown).

Although the degree of proliferation in the absence of factor shown by cells infected with the V449E, I374N and FIΔ mutants was lower than that seen with factor, this probably reflects the fact that only a proportion of each G418-selected population expressed hβc (Figure 3). To eliminate the possibility that the factor-independent populations represented secondary mutations that conferred a selective advantage, we performed colony assays on FDC-P1 cells, immediately after infection, in the presence or absence of mGM-CSF. The results (Table 1) show that between 28 and 78% of clonogenic cells infected with viruses encoding V449E, I374N or FIΔ were factor-independent; this proportion is far too high to be due to secondary mutations in a subset of infected cells. When these same G418-resistant populations were subsequently selected for factor independence, the proliferation observed in the absence of factor was comparable with that of uninfected FDC-P1 cells in the presence of mGM-CSF (Figure 4B). The presence of the appropriate hβc cDNA constructs in the factor-independent FDC-P1 cells was further confirmed by recovery of the entire hβc fragment by PCR from genomic DNA, followed by restriction enzyme digestions diagnostic of each mutant (Figure 7B). We therefore concluded that the V449E and I374N substitutions could confer factor independence on FDC-P1 cells and thus could constitutively activate hβc. Note that the latter mutation results in the loss of a *Bst*YI restriction site (as shown in Figure 7); by digesting with *Bst*YI, we subsequently found that all nine of the factor-independent clones that lacked the V449E mutation carried the I374N mutation (data not shown).

Biological activity of hβc mutants expressed in other cell types

While the FDC-P1 cells used in the preceding studies do not express human α subunits, they do express murine GMRα and IL-3Rα subunits. Because signalling by normal GM-CSF, IL-3 and IL-5 receptors requires formation of a complex of α subunits, β subunits and ligand, we wished to investigate whether the ligand-independent mutants could function in the absence of any α subunits. For these studies, we chose the murine IL-2-dependent T-cell line CTLL-2 (Cerottini *et al.*, 1974) because it does not express

Table I. Frequency of factor independence following infection of FDC-P1 cells with hβc mutants[a]

FDC-P1 cells infected with[b]:	Number of colonies[c]			
	mGM-CSF	mGM-CSF + G418	No factor	Percent factor-independent[d]
No virus	550	0	0	0
Wild-type hβc	689	562	0	0
I374N	936	736	573	77.9
V449E	886	565	157	27.8
FIΔ	587	491	305	62.1

[a]Cells were washed and plated in agar-containing medium, as described in Materials and methods, immediately after co-cultivation with Ψ2 cells.
[b]Virus-producing cells were sorted by flow cytometry for those expressing hβc prior to use in this experiment.
[c]Average number of colonies present on duplicate agar plates seeded with 10^3 cells.
[d]Calculated as the percentage of infected, i.e. G418-resistant, colonies that were factor-independent.

any endogenous GM-CSF/IL-3/IL-5 receptor chains but has been shown to proliferate in the presence of human GM-CSF following introduction of both receptor subunits (Kitamura *et al.*, 1991b).

Although infection of CTLL-2 cells by the murine retroviruses used was very inefficient (probably ≤ 0.01% as estimated by G418 resistance; unpublished observations) we were able to isolate CTLL-2 infectants, expressing both wild-type and mutant (I374N, V449E and FIΔ) forms of hβc, by selecting for G418 resistance and maintaining the cells in the presence of IL-2. Moreover, flow cytometric analysis (Figure 5A) showed that a substantial proportion of these cells clearly exhibited cell surface expression of β subunits. Somewhat to our surprise, however, the resulting cell lines failed to proliferate in the absence of IL-2 (Figure 6A). Because of the low frequency of infection, and the consequent possibility that the CTLL-2 populations may have comprised only one or a few clones, we wondered whether they were expressing aberrant, i.e. spontaneously mutated, forms of the original constructs. However, recovery of the hβc cDNAs from these cells by PCR generated products of the expected size and which exhibited the expected restriction enzyme digestion patterns (Figure 7C).

One possible explanation for this result was that the activated hβc mutants function by interacting with an α subunit (see also Discussion) and thus were unable to function in CTLL-2 cells lacking α subunits, but could interact with the GMRα or IL-3Rα subunits in the FDC-P1 cells. To test this possibility, and also to confirm that the CTLL-2 lines we derived could in fact respond to normal GMR-generated signals, we superinfected CTLL-2 cells expressing either wild-type or mutant β subunits with a vector carrying the hGMRα subunit and a puromycin resistance gene (pRUFPuro-hGMRα). We reasoned that if the mutant β subunits interacted with mouse α subunits, then it would be likely that they would also interact with human α subunits.

We were able to derive puromycin-resistant populations from each β subunit-expressing CTLL-2 line (i.e. wild-type, I374N, V449E and FIΔ); as only a proportion of these (12.3–33.2%) expressed the α subunit (data not shown), the latter were purified by preparative flow cytometry. Re-analysis of the sorted cells showed that they expressed both subunits of the human GMR (Figure 5B). Nevertheless, none of the resulting populations could proliferate in the absence of IL-2 (Figure 6B), indicating that the presence of the hGMRα subunit did not enable

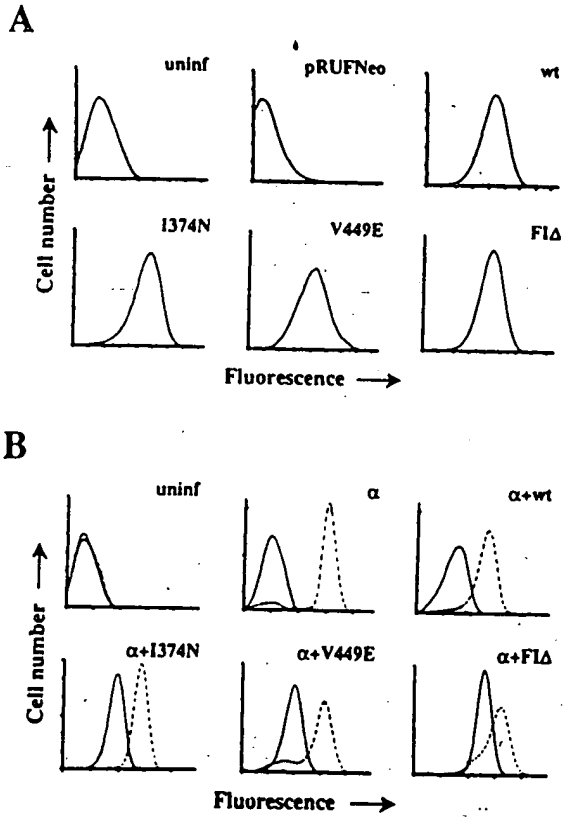


Fig. 5. Flow cytometric analysis of hβc and hGMRα expression on infected CTLL-2 cells. Nomenclature and axes are as in Figure 3. Dotted lines (.....) represent cells stained with an irrelevant control antibody, solid lines (——) indicate staining with anti-hβc antibody and dashed lines (-------) indicate staining with anti-hGMRα antibody. (A) CTLL-2 cells infected with retroviruses encoding the hβc mutants indicated on each histogram were stained with an anti-hβc monoclonal antibody as described in Materials and methods. (B) CTLL-2 cells infected with retroviruses encoding the hβc mutants indicated on each histogram were superinfected with a retrovirus encoding hGMRα and stained with anti-hβc and anti-hGMRα monoclonal antibodies. For comparison, analyses of uninfected CTLL-2 cells and cells infected only with the hGMRα virus (α) are also shown.

the constitutive generation of proliferative signals by either hβc mutant.

The possibility remained, though, that despite the PCR analysis (Figure 7C) there were other defects in the β subunits expressed by the CTLL-2 cells or that our CTLL-2 cells could not respond to GMR-generated signals. We therefore transferred CTLL-2 cells co-expressing α and β subunits from medium containing IL-2 to medium con-

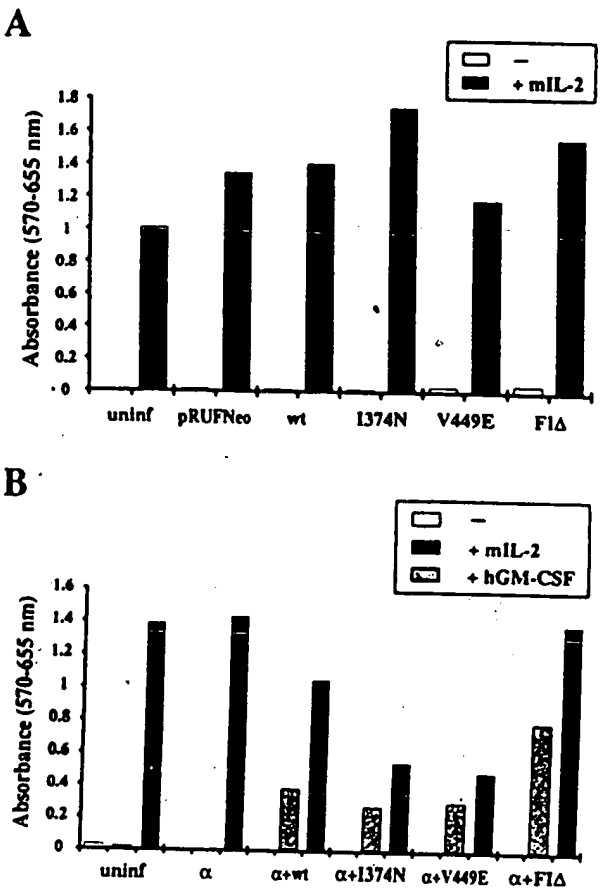


Fig. 6. Proliferation of CTLL-2 cells infected with hβc mutants in the presence and absence of human GM-CSF or mouse IL-2.
(A) Proliferation of cells selected and maintained prior to assay in IL-2 plus (except for the uninfected control cells) G418. Proliferation assays were carried out, as described in Materials and methods, in the presence or absence of IL-2 for uninfected CTLL-2 cells (uninf), cells infected with the vector alone (pRUFNeo), wild-type hβc (wt) or the indicated hβc mutants. (B) Proliferation of CTLL-2 cells infected with activated hβc mutants and superinfected with a retrovirus encoding hGMRα. Proliferation assays were carried out in the presence of mIL-2, hGM-CSF or in the absence of either factor, as indicated. Proliferation of uninfected CTLL-2 cells (uninf) and cells infected only with the hGMRα (α) virus is shown for comparison.

taining hGM-CSF. Cells expressing the wild-type hβc, and also those expressing the mutant β subunits, continued to grow in hGM-CSF, as shown by proliferation assays (Figure 6B); in fact, growth could be maintained in as little as 1 ng/ml hGM-CSF (data not shown) which would be expected to stimulate only high-affinity receptors. This indicates that the mutant receptors could still interact in an apparently normal manner with ligand and α subunits.

In view of these results, we wondered whether the ability of the mutants to confer factor independence was restricted to FDC-P1 cells. Retroviruses containing the mutant and wild-type forms of hβc were therefore used to infect the IL-3-dependent BAF-B03 subline of the pro-B cell line, Ba/F3. Figure 8A shows that, as before, surface expression of wild-type hβc and each mutant (V449E, I374N and FIΔ) could be readily obtained. However, only V449E could confer factor independence on these cells, as shown by proliferation assays (Figure 8B) and prolonged monitoring of liquid cultures in factor-free medium (data not shown). As seen with the CTLL-2 cells, superinfection f BAF-F03 cells expressing I374N

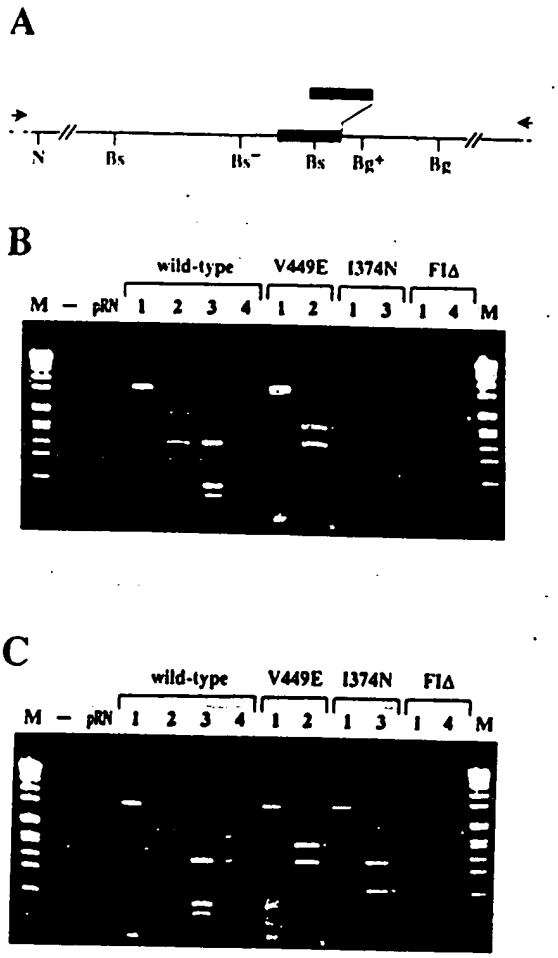


Fig. 7. Analysis of proviral hβc sequences in infected cells by PCR and restriction enzyme digestion. (A) Map of hβc cDNA showing *Nco*I (N), *Bst*YI (Bs) and *Bgl*II (Bg) restriction sites used to authenticate each form of hβc, as well as the region duplicated in FIΔ (indicated by boxes). The restriction sites affected by the point mutations are indicated as Bg^+ (gained in V449E) and Bs^- (lost in I374N). Arrows indicate the positions of PCR primers used to amplify hβc fragments from genomic DNA. (B) Electrophoretic analysis of PCR products generated from genomic DNA of FDC-P1 cells infected with the indicated hβc mutants; controls are reactions containing no DNA (–) or DNA from cells infected with the vector alone (pRN). Lanes M contain DNA size standards [SPP-1 phage DNA digested with *Eco*RI (Bresatec Ltd, Adelaide, South Australia)]. PCR products were either undigested (lanes 1), digested with *Bgl*II (lanes 2), *Bst*YI (lanes 3) or *Nco*I plus *Bgl*II (lanes 4). Bands in each digest that differ between the mutants and the wild-type are indicated by asterisks in the lanes containing the DNA from the appropriate mutant-infected cells. (C) Electrophoretic analysis, as in (B), of PCR products generated from genomic DNA of CTLL-2 cells infected with the indicated hβc mutants.

and FIΔ with the hGMRα subunit failed to confer factor independence (data not shown). Nevertheless, hGMRα subunit expression did allow these cells to proliferate in 1 ng/ml hGM-CSF (data not shown).

Tumorigenicity of hβc point mutants

The ability to confer factor independence on haemopoietic cell lines is generally indicative of tumorigenic potential (see Discussion). To test whether the V449E and I374N mutations were capable of conferring tumorigenicity, we injected factor-independent FDC-P1 cells expressing each of the mutant β subunits into syngeneic mice. Uninfected

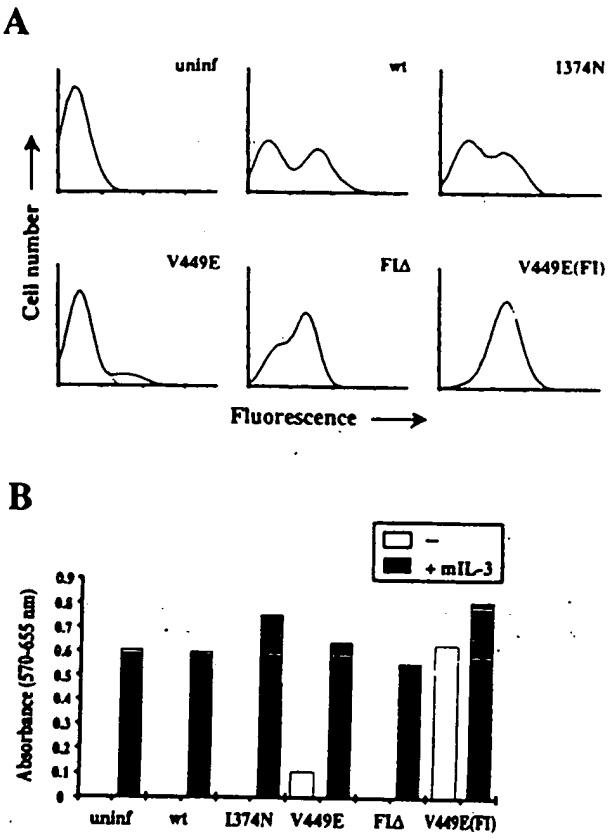


Fig. 8. Analysis of BAF-F03 cells infected with retroviruses encoding activated forms of hβc. (A) Flow cytometric analysis of hβc expression. Procedures and nomenclature are as for Figure 3, except that the panel labelled 'V449E(FI)' shows staining of BAF-F03 cells infected with the V449E mutant and selected for factor-independent growth. (B) Proliferation of the BAF-F03 cells depicted in (A) in the presence and absence of mouse IL-3.

FDC-P1 cells and FDC-P1 cells expressing wild-type hβc were injected as controls. All four mice injected with V449E, I374N and, as expected (D'Andrea *et al.*, 1994), FIΔ-expressing cells developed tumours at the site of injection within 3 weeks, while none of the control mice had developed tumours by 9 weeks post-inoculation.

Discussion

Oncogenic potential of the GM-CSF receptor

In this report, we have shown that hβc can be activated by point mutations leading to single amino acid substitutions. To date, the only other member of the cytokine receptor family shown to be activated by point mutation is the erythropoietin receptor; in this case activation resulted from the substitution of cysteine residues in the extracelluar domain (Yoshimura *et al.*, 1990; Watowich *et al.*, 1994). It seems quite likely that there may also be other point mutations that can activate hβc. For example, the studies described here have targeted only about one-third of the hβc cDNA, so there may well be potential activating mutations in other parts of the molecule and/or in associated α subunits.

These findings, and the previous report of a structural rearrangement that leads to activation of hβc (i.e. the FIΔ mutant; D'Andrea *et al.*, 1994), raise the possibility that activating mutations in hβc may c ntribute to human malignancies—most likely, myeloid leukaemias. Lesions that confer factor independence on FDC-P1 and similar factor-dependent cell lines, such as autocrine GM-CSF production (Lang *et al.*, 1985), expression of activated *abl* (Cook *et al.*, 1985) and *fms* (Wheeler *et al.*, 1987) genes, usually lead to the acquisition of tumorigenicity. This is also true for the activating mutations in hβc, as described here and by D'Andrea *et al.* (1994). Furthermore, many of these same lesions have also been implicated in myeloid leukaemias (e.g. Shtivelman *et al.*, 1985; Young and Griffin, 1986; Ridge *et al.*, 1990, respectively). The only survey (that we are aware of) which has examined the hβc gene in human leukaemias revealed no evidence for rearrangements (Brown *et al.*, 1993). However, that study would not have detected point mutations, as it was performed by Southern analysis of genomic DNA. Extending our present approach to provide a comprehensive 'map' of potential activating mutations in hβc will provide a rational basis for selecting, for finer analysis, particular regions of the hβc gene in leukaemias.

Implications for the mechanism of action of normal and mutant GM-CSF receptors

Understanding the mechanism by which the mutations result in receptor activation will be intimately connected to understanding how the normal GMR is triggered by GM-CSF. Both of the main findings of this study—the nature of the activating mutations and the differing activities of the mutants in a range of cell types—bear on this issue. The differing abilities of the mutants to confer factor independence on BAF-F03 cells indicate that they exert their effects by distinct mechanisms, which is consistent with the positions of the mutations within hβc. Furthermore, our results imply that the mutants, and, most likely, the normal receptors, interact with cell-type specific molecules. It is difficult to avoid the conclusion that there are (at least) two such molecules present in FDC-P1 cells that are involved: one, which is required for signalling by the V449E mutant, is also present in BAF-F03 cells but is absent from CTLL-2 cells; and a second which is required for signalling by the I374N and FIΔ mutants but which is absent from both BAF-F03 and CTLL-2 cells. It is unlikely that either of these components are α subunits, as co-expressing the human GMRα subunit (Figure 6) or either the human or murine IL-3Rα subunits (data not shown) with the hβc mutants in CTLL-2 cells did not result in factor-independent growth. Similarly, introduction of the human α subunits failed to allow BAF-F03 cells expressing the I374N or FIΔ mutants to proliferate in the absence of factor (data not shown).

Mechanism of activation by the V449E mutation. At the moment, the transmembrane domain mutation V449E is probably the simplest to interpret since it is analogous to the activating mutation (V664E) found in the rat *neu* oncogene (Bargmann *et al.*, 1986). Several reports have indicated that the *neu* mutation probably acts by inducing constitutive receptor dimerization (Sternberg and Gullick, 1989; Weiner *et al.*, 1989). Since dimerization is believed to be the key step in ligand-induced signalling by all receptor tyrosine kinases (reviewed by Ullrich and Schlessinger, 1990), the mutation has essentially the same effect as ligand binding. By analogy, we propose that the

V449E mutation acts by inducing ligand-independent dimerization of the β subunit and initiating the generation of intracellular signals.

The exact basis for dimerization induced by hydrophilic substitutions in the transmembrane domain is unclear, even though models have been proposed for Neu that invoke particular interactions between residues in the mutant transmembrane domains (Sternberg and Gullick, 1989) or conformational alterations (Brandt-Rauf *et al.*, 1990). Our model for V449E is based on the striking similarity to activating mutations in Neu. However, the sequence surrounding the mutated valine in hβc does not conform to the 'rules' worked out for Neu (Cao *et al.*, 1992), although both hβc and Neu have an alanine four residues N-terminal to the substituted valine. Moreover, not all receptors can be activated by such mutations. Thus it is very difficult to speculate as to why the V449Q and V449D mutations do not activate hβc; structural or modelling data will be necessary to provide further insight into the requirements for activation.

The notion that β subunit dimerization can trigger intracellular signalling is supported by a number of observations pertaining to murine or human β subunits and to other members of the cytokine receptor family. Several members of this family (e.g. the receptors for G-CSF, Epo and human growth hormone) clearly function as homodimers; in these cases ligand binding induces dimerization which, in turn, results in the generation of proliferative signals (Fukunaga *et al.*, 1990; de Vos *et al.*, 1992; Watowich *et al.*, 1994). A second important case is that of the IL-6 receptor which, like the GMR, is heteromeric (Taga *et al.*, 1989) and comprises a ligand binding α subunit plus the signalling subunit, gp130 (which can be viewed as the equivalent of the GMR β subunit). Activation of the IL-6 receptor by ligand results in the formation of a heteromeric complex that contains a gp130 dimer (Murakami *et al.*, 1993) along with two IL-6 receptor α subunits (Ward *et al.*, 1994). With regard to GMR/IL-3R/IL-5R β subunits, it has been shown that a chimera containing the extracellular region of the Epo receptor and the intracellular domain of the murine $\beta_{IL\text{-}3}$ (AIC2A) subunit can mediate Epo-dependent proliferation (at least in Ba/F3 cells; Sakamaki *et al.*, 1993). In addition, other experiments have shown that chimeras comprising the extracellular region of the GMRα subunit (K.Arai, personal communication) or the IL-5Rα subunit (Takaki *et al.*, 1994) and the intracellular region of hβc can transduce proliferative signals in the presence of normal hβc and the appropriate ligand. These data imply that dimerization of the intracellular portion of the β subunit is sufficient to initiate cellular proliferation.

One model that could explain the inability of the V449E mutant (see Figure 6) and the EpoR–$\beta_{IL\text{-}3}$ chimera (Sakamaki *et al.*, 1993) to function in CTLL-2 cells is illustrated in Figure 9A. The model assumes that the normal receptor complex (or at least one form thereof—see below) comprises two α and two β subunits, and proposes: (i) that there is a factor present in, for example, FDC-P1 cells (factor 'Y') which must be associated with hβc homodimers in order to generate proliferative signals but which is absent from CTLL-2 cells; and (ii) that association of α subunit plus ligand can compensate for the lack of factor Y by bringing into the complex another factor ('X', which is present in CTLL-2 cells and probably in other cells also) which overlaps in function with factor Y. Thus, the dimerized β subunit could deliver proliferative signals in cells such as FDC-P1 and BAF-F03 where Y is present, but could only signal in CTLL-2 cells following ligand-induced association with an α subunit.

Possible candidates for molecules X and Y might be non-receptor tyrosine kinases such as members of the *src* family, as the IL-3R and the GMR have been reported to associate with and/or activate the products of the *fes* (Hanazono *et al.*, 1993) and the *src*-related *lyn* (Torigoe *et al.*, 1992) genes. While the JAK-2 kinase is also associated with the IL-3R (Silvennoinen *et al.*, 1993), it is unlikely that X or Y is JAK-2 since the latter is also present in CTLL-2 cells (Mano *et al.*, 1993); it is more likely that JAK kinases are required in addition to X or Y. Alternatively, X and Y may be adaptor molecules that, by analogy with GRB2 (Lowenstein *et al.*, 1992) for example, bring other components of the signal transduction pathway into the receptor complex. A third possibility is that they are members of the STAT family of transcription factors that associate with many members of the cytokine receptor family (reviewed by Darnell *et al.*, 1994).

Mechanism of activation by the I374N and FIΔ mutations. For reasons discussed above—in particular their differing abilities to function in BAF-F03 cells—these mutations are unlikely to act in the same way as V449E, i.e. by formation of hβc homodimers. Moreover, as both are within the extracellular portion of hβc, it appears likely that they may function by interacting with another membrane-spanning molecule. Such a situation would be superficially reminiscent of the case of the IL-2 receptor which has a third component, the γ subunit (Takeshita *et al.*, 1992). However, the analogy is not strictly appropriate because the IL-2R γ subunit appears to be essential for signalling (reviewed by Taniguchi and Minami, 1993). One could, however, postulate that a myeloid-specific 'γ subunit' could associate with an αβ dimer to provide a function which overlaps that of a second β subunit. Functional receptors could then have one of two compositions—$(\alpha\beta)_2$ or αβγ. Thus the I374N and FIΔ mutants may form βγ complexes constitutively, i.e. in the absence of α subunits or ligand, and generate proliferative signals in myeloid (e.g. FDC-P1) cells but not in CTLL-2 or BAF-F03 cells, which represent the T- and B-cell lineages, respectively. Consistent with this, preliminary studies have indicated that FIΔ (and V449E) can induce factor-independent colony formation by myeloid progenitors from murine fetal liver (unpublished observations of M.McCormack and the authors). This model is illustrated in Figure 9B.

It is not clear how the I374N and FIΔ mutations could induce constitutive association with an additional, membrane-spanning receptor component such as a putative γ subunit. However, one plausible explanation comes from examining the structure of the regions involved. Sequence alignment of the extracellular regions of hβc and hGHR (Bazan, 1990; Goodall *et al.*, 1993) suggests that Ile374 lies in the β strand C of the membrane-proximal domain of hβc. Furthermore, predictions based on the structure of hGHR (de Vos *et al.*, 1992; C.J.Bagley and A.Lopez, personal communication) suggest that the side chain of

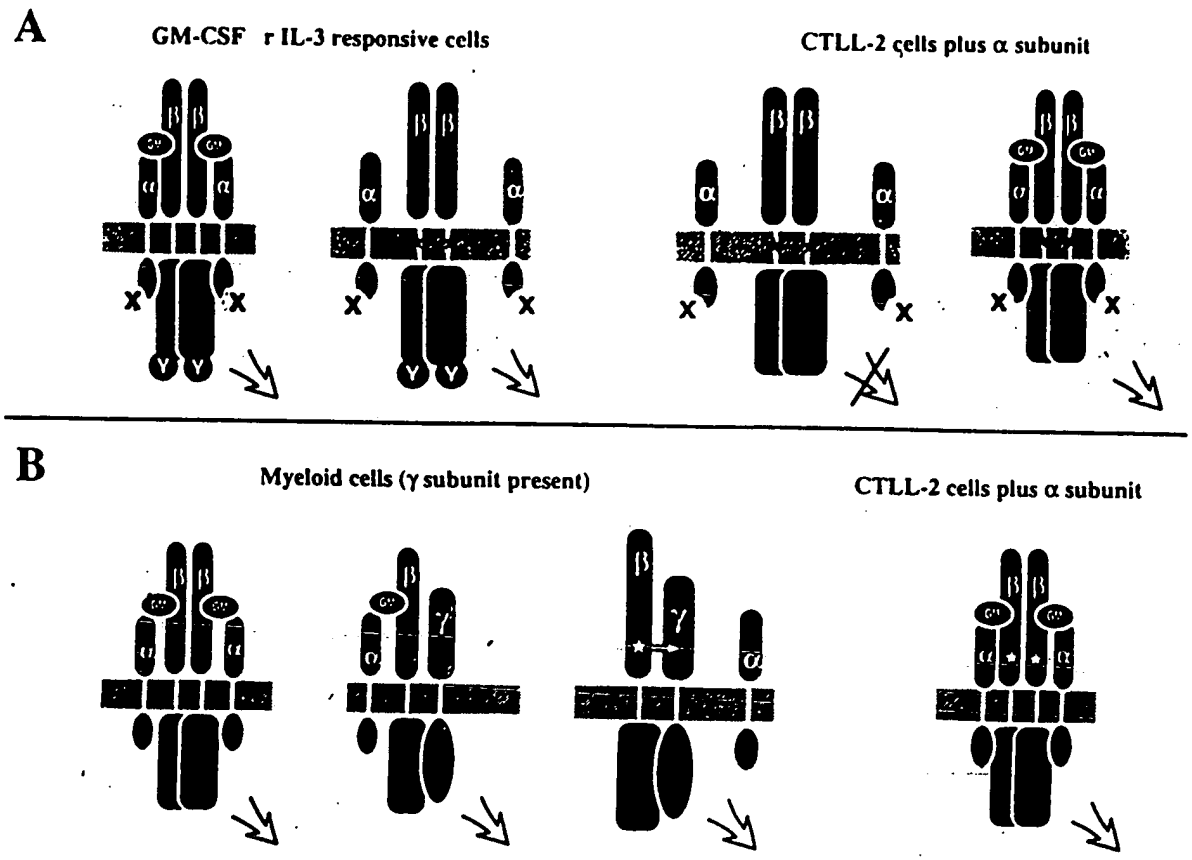


Fig. 9. Models for the mechanism of activation and cell-type specificity of constitutive hβc mutants. In each panel, the cell membrane is represented by the hatched rectangle, with the extracellular domains of the α, β and putative γ subunits shown above the rectangle. The positions of activating mutations are indicated by stars. The ligand indicated is GM-CSF (GM) but the models pertain equally to IL-3 and IL-5. Generation of proliferative signals is indicated by the large arrow at the bottom right of each panel. Both models assume that (one form of) the normal receptor complex contains two α plus two β subunits and that dimerization of the β subunit can trigger proliferative signalling. (A) A model pertaining to the V449E mutant which postulates that the transmembrane mutation results in ligand-independent hβc dimerization. This model also invokes the existence of an intracellular signalling molecule, Y, which is specific to IL-3- or GM-CSF-responsive cells (i.e. FDC-P1 or BAF-F03 in this study) and which associates with the β subunit. Molecule X, which can substitute for Y in the receptor complex, associates with the α subunit. Proliferative signalling requires a complex containing either X or Y. See text for further explanation. (B) A model pertaining to the I374N and FIΔ mutants which invokes the existence of a third receptor subunit, a 'γ' subunit, which is specific for myeloid cells. The model postulates that the γ subunit can associate with the β subunit of an αβ heterodimer and can contribute to proliferative signalling in concert with the β subunit. Thus, two alternative forms of the normal receptor complex, $(\alpha\beta)_2$ or $\alpha\beta\gamma$, can exist in myeloid cells; the I374N and FIΔ mutations would then function by allowing constitutive association with the γ subunit. In both (A) and (B) the mutant β subunits can form functional $(\alpha\beta)_2$ complexes in CTLL-2 cells when both α subunit and ligand are present.

Ile374 may contact the hydrophobic side chains of valine residues in the conserved RVRVR sequence within the F strand of hβc. Intriguingly, the F strand is part of the region duplicated in the FIΔ mutant; thus, both the duplication in the FIΔ mutant and the hydrophilic substitution in I374N may be affecting a common structure or surface that includes strands F and C—and which, we suggest, may be involved in interacting with the putative γ subunit. A recent report by Grube and Cochrane (1994) supports the notion that strand C may play a regulatory role in interactions with other receptor subunits, as it was found that a peptide corresponding to strand C of the IL-6 receptor (α subunit) can inhibit signalling by IL-6 without affecting ligand binding. We therefore suggest, in a similar vein to D'Andrea *et al.* (1994), that association with ligand plus α subunit, a disruptive amino acid substitution such as I374N in the β strand C or duplication of strand F may all act to unmask an interactive surface involving strand F, allowing association f hβc with the putative γ subunit and triggering the generation of intracellular signals.

Both of the models illustrated in Figure 9 require a degree of functional redundancy between components of the GMR (i.e. between X and Y, and between the β and 'γ' subunits). This redundancy may not be complete, however, because the studies presented here only address proliferation in established cell lines; it is possible that the 'redundant' components may vary in their abilities to generate, say, differentiative signals.

We also wish to point out the potentially general utility of the methodology used to isolate the mutants described here. While a retroviral library of random mutants has already been used by Druker and Roberts (1991), the PCR mutagenesis approach used here is much simpler than the chemical mutagenesis employed by these authors. Because many receptors, including members of both receptor tyrosine kinase and cytokine receptor families, can induce ligand-dependent proliferation of factor-dependent haemopoietic cell lines (e.g. Di Fiore *et al.*, 1990; D'Andrea *et al.*, 1991; Sakamaki *et al.*, 1993; Ziegler *et al.*, 1993), the procedure described here offers the possibility of identifying activating mutations in any of these molecules.

Furthermore, this option may als apply to intracellular molecules that are involved in receptor-mediated signalling. Finally, it should be possible to extend this approach to ther types of screens using different cell types, including primary cells, and t devise screens for loss-of-function mutants as well as gain-of-function mutants. This could facilitate the identification of functional domains and structures in molecules involved in other cellular functions such as differentiation, adhesion or gene regulation.

Materials and methods

Cell lines and cDNAs

PA317 (Miller and Buttimore, 1986) and Ψ2 (Mann *et al.*, 1983) retrovirus packaging cell lines were maintained in DME medium supplemented with 10% fetal calf serum (FCS). The IL-3/GM-CSF-dependent mouse myeloid cell line, FDC-P1 (Dexter *et al.*, 1980), and the BAF-B03 subline (Hatakeyama *et al.*, 1989) of the IL-3-dependent mouse pro-B cell line, Ba/F3, were maintained in the same medium, as above, containing 80 units/ml mouse GM-CSF (produced by an engineered yeast strain and kindly supplied by Dr Tracy Wilson, Walter and Eliza Hall Institute, Melbourne) or 300 units/ml mouse IL-3 (produced from a baculovirus vector and kindly supplied by Dr Andrew Hapel, John Curtin School of Medical Research, Canberra), respectively. The IL-2-dependent mouse T-cell line, CTLL-2 (Cerottini *et al.*, 1974), was maintained in RPMI-1640 medium supplemented with 10% FCS and 50 μM 2-mercaptoethanol, in the presence of 100 units/ml bacterially synthesized mouse IL-2 (expressed from the plasmid pTRC11 which was kindly provided by Dr G.Zurawski, DNAX, Palo Alto) or 20% conditioned medium from MLA cells (CTLL medium).

The hβc cDNA (Hayashida *et al.*, 1990) used here was that described by Barry *et al.* (1994). cDNA for hGMRα (Gearing *et al.*, 1989) was kindly provided by Dr Nic Nicola (Walter and Eliza Hall Institute, Melbourne, Australia). The activated FIΔ mutant of hβc has been described previously (D'Andrea *et al.*, 1994).

Site-directed mutagenesis and construction of expression plasmids

Site-directed mutagenesis was carried out on single-stranded DNA with mutagenic oligonucleotides using the pAlter-1 system (Promega) in accordance with the manufacturer's instructions. All mutations were confirmed by DNA sequencing, following which mutant hβc cDNAs were blunt-end ligated between end-filled *Bgl*II and *Xho*I restriction sites of the pRUFNeo retroviral vector. The pRUFNeo retroviral vector was modified from the version described by Rayner and Gonda (1994) by the removal of the *Nco*I restriction site from the multiple cloning site.

The pRUFPuro retroviral expression vector was constructed by replacing the MC1Neo cassette of pRUFNeo with an SV40/puromycin resistance cassette from pBabePuro (Morgenstern and Land, 1990). A vector for expressing hGMRα was constructed by inserting the hGMRα cDNA into the *Xho*I site of pRUFPuro.

To facilitate cloning of hβc PCR fragments (see below), a silent mutation creating a *Xho*I restriction site was introduced into the wild-type hβc cDNA by changing nucleotide 772 (Hayashida *et al.*, 1990; sequence accession number M38275) from G to C. This cDNA was then inserted into pRUFNeo as above to generate pRUFNeo-hβc.

PCR mutagenesis and construction of a point-mutated hβc cDNA library

Random point mutations were introduced into a 940 bp *Xho*I–*Bgl*II segment, bases 770–1710 (Hayashida *et al.*, 1990) of the hβc cDNA, by PCR mutagenesis based on the method described by Cadwell and Joyce (1992). The primers used for amplification were 5'-TGGAGCCCAG-AGGTTTGCTGGGACT-3' (nucleotides 713–737) and 5'-GGGCCC-ATTGAAGTCAAAGCTGGAA-3' (nucleotides 1804–1780), defining a 1091 bp fragment. Standard and mutagenic reactions (in a volume of 50 μl) were performed on 5 ng of pRUFNeo-hβc plasmid DNA using 12 μM of each primer, and were cycled in a MiniCycler (MJ Research) for 30 cycles of 94°C for 1 min, 68°C for 1 min and 72°C for 1 min. Reactions were denatured at 95°C for 5 min prior to cycling. Standard reactions contained an equimolar concentration of dNTPs (0.2 mM), 1.5 mM $MgCl_2$ and 2.5 units of *Taq* polymerase (Perkin Elmer), whereas mutagenic reactions contained 0.2 mM dATP and dGTP, 1 mM dCTP and dTTP, .7 mM $MgCl_2$, 5 units of *Taq* polymerase and varying concentrations of $MnCl_2$. To estimate mutation frequencies, PCR products were purified from agarose gels, subcloned into appropriate pGEM vectors (Promega) and sequenced; the results are shown in Table II. No strong mutational bias was observed under these reaction conditions, although there was a weak bias towards T to N and N to A substitutions similar to that reported by Cadwell and Joyce (1992).

Table II. Effect of Mg^{2+} and Mn^{2+} concentrations on PCR mutation rate

[$MnCl_2$] mM	[$MgCl_2$] mM	Bases sequenced	Mutation rate
0	7.0	2520	1/840
0.1	7.0	3900	1/300
0.25	7.0	3060	1/510
0.5	7.0	2800	1/400
0	1.5	2200	1/2200

A library of hβc fragments containing mutations at the desired frequency (0.2%) was regenerated by performing the mutagenic PCR as above in the presence of 0.25 mM $MnCl_2$. Following digestion with *Xho*I and *Bgl*II, mutant 940 bp fragments were agarose gel-purified and ligated directionally into the pRUFNeo-hβc from which the *Xho*I–*Bgl*II fragment of hβc had been excised. Following transformation of *Escherichia coli* (DH10B), the resultant hβc cDNA point-mutant library (pRUFNeo-$hβc^{mut}$) of 1.7×10^5 independent plasmid clones was further amplified as described previously (Rayner and Gonda, 1994).

Expression of mutant hβc cDNAs in factor-independent FDC-P1 cells

Retroviral DNA was used to generate a library of retroviruses as described by Rayner and Gonda (1994). Briefly, PA317 amphotropic packaging cells were transfected with pRUFNeo-$hβc^{mut}$ plasmid DNA and virus-containing transient supernatants were filtered and used to infect Ψ2 ecotropic packaging cells. Infected Ψ2 cells were selected in medium containing 400 μg/ml G418, pooled, and used to infect FDC-P1 cells. This was achieved by co-cultivating 5×10^5 FDC-P1 cells with 10^6 irradiated (30 Gy) Ψ2 producer cells for 48 h in each of four 25 cm^2 flasks. The FDC-P1 cells from each flask were harvested, washed and selected for factor-independent growth in liquid culture medium containing G418 (1 mg/ml) without GM-CSF. Plating in soft agar was performed essentially as described by Johnson (1980); GM-CSF (80 units/ml) or G418 (1 mg/ml) was added as required.

Rescue of mutant hβc cDNAs from factor-independent cells

Genomic DNA was isolated from cells using a Proteinase K/SDS procedure essentially as described by Hughes *et al.* (1979). PCR was performed on 100 ng of genomic DNA with *Pfu* DNA polymerase (Stratagene) under conditions recommended by the manufacturer. The primers and cycling parameters used were as described above. PCR products were inserted into pGEM vectors (Promega), and sequenced.

Characterization of activating hβc point mutations

Point mutations generated in hβc by site-directed mutagenesis were tested for their ability to confer factor independence by expression in FDC-P1 cells or CTLL-2 cells. Ψ2 cells were transfected by the calcium phosphate procedure (as described by Rayner and Gonda, 1994) and stable transfectants selected with G418 (400 μg/ml) prior to co-cultivation. FDC-P1 and BAF-F03 cells were infected by co-cultivation and selected for factor-independent growth as before. For CTLL-2 infections, 4×10^6 stably transfected Ψ2 cells were irradiated (30 Gy) and co-cultivated with 10^6 CTLL-2 cells in CTLL medium (see above) for 2 days in 75 cm^2 flasks. CTLL-2 cells were separated from the adherent Ψ2 cells and selected in CTLL medium containing G418 at 750 μg/ml. After several weeks in culture, selected cells were washed and incubated in medium with or without factor.

To generate CTLL-2 cells co-expressing wild-type hGMRα and hβc mutants, Ψ2 cells were transfected with pRUFPuro-hGMRα plasmid DNA and stable transfectants selected with puromycin at 2 μg/ml. CTLL-2 cells expressing hβc mutants were infected by co-cultivation as before and selected in CTLL medium containing puromycin at 2 μg/ml.

Cell sorting and analysis of receptor subunit expression by flow cytometry

Selected cells expressing hGMRα were collected by cell sorting on a FACStar^{PLUS} flow cytometer (Becton-Dickinson). Briefly, cells were washed and resuspended in cold RPMI-1640 medium supplemented with 5% FCS (RPMI-FCS). Cells were incubated with the anti-hGMRα monoclonal antibody 8G6 (Q.Sun and A.Lopez, manuscript in preparation) for 20 min on ice, washed and subsequently incubated with FITC-conjugated anti-mouse IgG (Silenus) for 20 min on ice. After washing and resuspension in cold RPMI-FCS, the cells were sorted, and the positive cell population collected in CTLL medium.

Expression of hβc mutants on the cell surface of infected FDC-P1, BAF-F03 or CTLL-2 cells was confirmed by flow cytometric analysis using an Epics-Profile II analyser (Coulter) by staining, as described above, with the anti-hβc monoclonal antibody 3D7 (Q.Sun and A.Lopez, manuscript in preparation). hGMRα expression was analysed as above by staining with the monoclonal antibody 8G6.

Cell proliferation assays

Infected cells (FDC-P1, BAF-F03 or CTLL-2) were washed twice and triplicate samples of equal cell number (5×10^3) were cultured in a 96-well microtitre plate with or without appropriate growth factor for 72 h. Cell proliferation was measured by the CellTiter'96 Non-Radioactive Cell Proliferation Assay (Promega).

Tumorigenicity studies

FDC-P1 cells (5×10^6) expressing each of the mutant β subunits (V449E, I374N or FIΔ) were injected subcutaneously into each of four 10-week-old male DBA2 mice; as controls, mice were injected with cells expressing wild-type hβc or uninfected FDC-P1 cells. Tumour formation was monitored at weekly intervals.

Acknowledgements

The authors are grateful to Sun Qiyu and Angel Lopez for supplying their anti-receptor antibodies, Nic Nicola for supplying us with the hGMRα cDNA, our various colleagues who generously supplied us with growth factors, Alan Bishop for assistance with cell sorting and Chris Bagley for valuable insights into receptor structure. We thank Angel Lopez, Greg Goodall, Simon Barry and Mathew Vadas for many valuable discussions and for critical reading of the manuscript. This work was supported by a research grant (to T.J.G.) from the Anti-Cancer Foundation of the Universities of South Australia. T.J.G. is a Research Fellow of the National Health and Medical Research Council of Australia.

References

Bargmann,C.I. and Weinberg,R.A. (1988) Oncogenic activation of the *neu*-encoded receptor protein by point mutation and deletion. *EMBO J.*, 7, 2043–2052.

Bargmann,C.I., Hung,M.-C. and Weinberg,R.A. (1986) Multiple independent activations of the *neu* oncogene by a point mutation altering the transmembrane domain of p185. *Cell*, 45, 649–657.

Barry,S.C. *et al.* (1994) Two contiguous residues in human interleukin-3, Asp^{21} and Glu^{22}, selectively interact with the α- and β-chains of its receptor and participate in function. *J. Biol. Chem.*, 269, 8488–8492.

Bazan,J.F. (1990) Structural design and molecular evolution of a cytokine receptor superfamily. *Proc. Natl Acad. Sci. USA*, 87, 6934–6938.

Brandt-Rauf,P.W., Rackovsky,S. and Pincus,M.R. (1990) Correlation of the structure of the transmembrane domain of the *neu* oncogene-encoded p185 protein with its function. *Proc. Natl Acad. Sci. USA*, 87, 8660–8664.

Brown,M.A., Gough,N.M., Willson,T.A., Rockman,S. and Begley,C.G. (1993) Structure and expression of the GM-CSF receptor α and β chain genes in human leukemia. *Leukemia*, 7, 63–74.

Cadwell,R.C. and Joyce,G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Applic.*, 2, 28–33.

Cao,H., Bangalore,L., Bormann,B.J. and Stern,D.F. (1992) A subdomain in the transmembrane domain is necessary for $p185^{neu*}$ activation. *EMBO J.*, 11, 923–932.

Cerottini,J.-C., Engers,H.D., MacDonald,H.R. and Brunner,K.T. (1974) Generation of cytotoxic T lymphocytes *in vitro*. *J. Exp. Med.*, 140, 703–717.

Clark,S.C. and Kamen,R. (1987) The human hematopoietic colony-stimulating factors. *Science*, 236, 1229–1237.

Cook,W.D., Metcalf,D., Nicola,N.A., Burgess,A.W. and Walker,F. (1985) Malignant transformation of a growth factor-dependent myeloid cell line by Abelson virus without evidence of an autocrine mechanism. *Cell*, 41, 677–683.

D'Andrea,A.D., Yoshimura,A., Youssoufian,H., Zon,L.I., Koo,J.-W. and Lodish,H.F. (1991) The cytoplasmic region of the erythropoietin receptor contains nonoverlapping positive and negative growth-regulatory domains. *Mol. Cell. Biol.*, 11, 1980–1987.

D'Andrea,R., Rayner,J., Moretti,P., Lopez,A., Goodall,G.J., Gonda,T.J. and Vadas,M. (1994) A mutation of the common receptor subunit for interleukin-3 (IL-3), granulocyte–macrophage colony-stimulating factor, and IL-5 that leads to ligand independence and tumorigenicity. *Blood*, 83, 2802–2808.

Darnell,J.E.Jr, Kerr,I.M. and Stark,G.R. (1994) Jak–STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, 264, 1415–1421.

De Vos,A.M., Ultsch,M. and Kossiakoff,A.A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, 255, 306–312.

Dexter,T.M., Garland,J., Scott,D., Scolnick,F. and Metcalf,D. (1980) Growth of factor-dependent hemopoietic precursor cell lines. *J. Exp. Med.*, 152, 1036–1047.

Di Fiore,P.P., Segatto,O., Taylor,W.G., Aaronson,S.A. and Pierce,J.H. (1990) EGF receptor and *erb*B-2 tyrosine kinase domains confer cell specificity for mitogenic signaling. *Science*, 248, 79–83.

Druker,B.J. and Roberts,T.M. (1991) Generation of a large library of point mutations in polyoma middle T antigen. *Nucleic Acids Res.*, 19, 6855–6861.

Duronio,V., Clark-Lewis,I., Federsppiel,B., Wieler,J.S. and Schrader,J.W. (1992) Tyrosine phosphorylation of receptor β subunits and common substrates in response to interleukin-3 and granulocyte–macrophage colony-stimulating factor. *J. Biol. Chem.*, 267, 21856–21863.

Fukunaga,R., Ishizaka-Ikeda,E. and Nagata,S. (1990) Purification and characterization of the receptor for murine granulocyte colony-stimulating factor. *J. Biol. Chem.*, 265, 14008–14015.

Gearing,D.P., King,J.A., Gough,N.M. and Nicola,N.A. (1989) Expression cloning of a receptor for human granulocyte–macrophage colony-stimulating factor. *EMBO J.*, 8, 3667–3676.

Goodall,G.J., Bagley,C.J., Vadas,M.A. and Lopez,A.F. (1993) A model for the interaction of the GM-CSF, IL-3 and IL-5 receptors with their ligands. *Growth Factors*, 8, 87–97.

Grube,B.J. and Cochrane,C.G. (1994) Identification of a regulatory domain of the interleukin-6 receptor. *J. Biol. Chem.*, 269, 20791–20797.

Hanazono,Y., Chiba,S., Sasaki,K., Mano,H., Miyajima,A., Arai,K., Yazaki,Y. and Hirai,H. (1993) c-*fps/fes* protein-tyrosine kinase is implicated in a signaling pathway triggered by granulocyte–macrophage colony-stimulating factor and interleukin-3. *EMBO J.*, 12, 1641–1646.

Hatakeyama,M., Mori,H., Doi,T. and Taniguchi,T. (1989) A restricted cytoplasmic region of IL-2 receptor β chain is essential for growth signal transduction but not for ligand binding and internalization. *Cell*, 59, 837–845.

Hayashida,K., Kitamura,T., Gorman,D.M., Arai,K., Yokota,T. and Miyajima,A. (1990) Molecular cloning of a second subunit of the receptor for human granulocyte–macrophage colony-stimulating factor (GM-CSF): reconstitution of a high-affinity GM-CSF receptor. *Proc. Natl Acad. Sci. USA*, 87, 9655–9659.

Hughes,S.H., Payvar,F., Spector,D., Schimke,R.T., Robinson,H.L., Payne,G.S., Bishop,J.M. and Varmus,H.E. (1979) Heterogeneity of genetic loci in chickens: analysis of endogenous viral and normal genes by cleavage of DNA with restriction endonucleases. *Cell*, 18, 347–359.

Johnson,G.R. (1980) Colony formation in agar by adult bone marrow multipotential hemopoietic cells. *J. Cell. Physiol.*, 103, 371–383.

Kitamura,T. and Miyajima,A. (1992) Functional reconstitution of the human interleukin-3 receptor. *Blood*, 80, 84–90.

Kitamura,T., Sato,N., Arai,K. and Miyajima,A. (1991a) Expression cloning of the human IL-3 receptor cDNA reveals a shared β subunit for the human IL-3 and GM-CSF receptors. *Cell*, 66, 1165–1174.

Kitamura,T., Hayashida,K., Sakamaki,K., Yokota,T., Arai,K. and Miyajima,A. (1991b) Reconstitution of functional receptors for human granulocyte/macrophage colony-stimulating factor (GM-CSF): evidence that the protein encoded by the AIC2B cDNA is a subunit of the murine GM-CSF receptor. *Proc. Natl Acad. Sci. USA*, 88, 5082–5086.

Lang,R.A., Metcalf,D., Gough,N.M., Dunn,A.R. and Gonda,T.J. (1985)

Expression of a hemopoietic growth factor cDNA in a factor-dependent cell line results in autonomous growth and tumorigenicity. *Cell*, 43, 531–542.

Lok,S. *et al.* (1994) Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production *in vivo*. *Nature*, 369, 565–568.

Lopez,A.F. *et al.* (1991) Interleukin-5, interleukin-3, and granulocyte–macrophage colony-stimulating factor cross-compete for binding to cell surface receptors on human eosinophils. *J. Biol. Chem.*, 266, 24741–24747.

Lowenstein,E.J. *et al.* (1992) The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell*, 70, 431–442.

Mann,R., Mulligan,R.C. and Baltimore,D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, 33, 153–159.

Mano,H., Mano,K., Tang,B., Koehler,M., Yi,T., Gilbert,D.J., Jenkins,N.A., Copeland,N.G. and Ihle,J.N. (1993) Expression of a novel form of *Tec* kinase in hematopoietic cells and mapping of the gene to chromosome 5 near *Kit*. *Oncogene*, 18, 417–424.

Metcalf,D. (1986) The molecular biology and functions of the granulocyte–macrophage colony-stimulating factors. *Blood*, 67, 257–267.

Miller,A.D. and Buttimore,C. (1986) Redesign of retrovirus packaging cell lines to avoid recombination leading to helper virus production. *Mol. Cell. Biol.*, 6, 2895–2902.

Miyajima,A., Kitamura,T., Harada,N., Yokota,T. and Arai,K. (1992) Cytokine receptors and signal transduction. *Annu. Rev. Immunol.*, 10, 295–331.

Morgenstern,J.P. and Land,H. (1990) Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res.*, 18, 3587–3596.

Murakami,M., Hibi,M., Nakagawa,N., Nakagawa,T., Yasukawa,K., Yamanishi,K., Taga,T. and Kishimoto,T. (1993) IL-6-induced homodimerization of gp130 and associated activation of a tyrosine kinase. *Science*, 260, 1808–1810.

Polotskaya,A., Zhao,Y., Lilly,M.B. and Kraft,A.S. (1994) Mapping the intracytoplasmic regions of the α granulocyte–macrophage colony-stimulating factor receptor necessary for cell growth regulation. *J. Biol. Chem.*, 269, 14607–14613.

Rayner,J.R. and Gonda,T.J. (1994) A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. *Mol. Cell. Biol.*, 14, 880–887.

Ridge,S.A., Worwood,M., Oscier,D., Jacobs,A. and Padua,R.A. (1990) *FMS* mutations in myelodysplastic, leukemic, and normal subjects. *Proc. Natl Acad. Sci. USA*, 87, 1377–1380.

Sakamaki,K., Miyajima,I., Kitamura,T. and Miyajima,A. (1992) Critical cytoplasmic domains of the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors for growth signal transduction and tyrosine phosphorylation. *EMBO J.*, 11, 3541–3549.

Sakamaki,K., Wang,H.-M., Miyajima,I., Kitamura,T., Todokoro,K., Harada,N. and Miyajima,A. (1993) Ligand-dependent activation of chimeric receptors with the cytoplasmic domain of the interleukin-3 receptor β subunit (β_{IL3}) *J. Biol. Chem.*, 268, 15833–15839.

Sanderson,C.J. (1992) Interleukin-5, eosinophils, and disease. *Blood*, 79, 3101–3109.

Sato,N., Sakamaki,K., Terada,N., Arai,K. and Miyajima,A. (1993) Signal transduction by the high-affinity GM-CSF receptor: two distinct cytoplasmic regions of the common β subunit responsible for different signalling. *EMBO J.*, 12, 4181–4189.

Shtivelman,E., Lifshitz,B., Gale,R.P. and Canaani,E. (1985) Fused transcript of *abl* and *bcr* genes in chronic myelogenous leukaemia. *Nature*, 315, 550–554.

Silvennoinen,O., Witthuhn,B.A., Quelle,F.W., Cleveland,J.L., Yi,T. and Ihle,J.N. (1993) Structure of the murine Jak2 protein-tyrosine kinase and its role in interleukin 3 signal transduction. *Proc. Natl Acad. Sci. USA*, 90, 8429–8433.

Souyri,M., Vigon,I., Penciolelli,J.-F., Heard,J.-M., Tambourin,P. and Wendling,F. (1990) A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalizes hematopoietic progenitors. *Cell*, 63, 1137–1147.

Sternberg,M.J.E. and Gullick,W.J. (1989) *Neu* receptor dimerization. *Nature*, 339, 587.

Taga,T., Hibi,M., Hirata,Y., Yamasaki,K., Yasukawa,K., Matsuda,T., Hirano,T. and Kishimoto,T. (1989) Interleukin-6 triggers the association of its receptor with a possible signal transducer, gp130. *Cell*, 58, 573–581.

Takaki,S., Murata,Y., Kitamura,T., Miyajima,A., Tominaga,A. and Takatsu,K. (1993) Reconstitution of the functional receptors for murine and human interleukin-5. *J. Exp. Med.*, 177, 1523–1529.

Takaki,S., Kanazawa,H., Shiba,M. and Takatsu,K. (1994) A critical cytoplasmic domain of the interleukin-5 (IL-5) receptor α chain and its function in IL-5-mediated growth signal transduction. *Mol. Cell. Biol.*, 14, 7404–7413.

Takeshita,T., Asao,H., Ohtani,K., Ishii,N., Kumaki,S., Tanaka,N., Munakata,H., Nakamura,M. and Sugamura,K. (1992) Cloning of the γ chain of the human IL-2 receptor. *Science*, 257, 379–382.

Taniguchi,T. and Minami,Y. (1993) The IL-2/IL-2 receptor system: a current overview. *Cell*, 73, 5–8.

Tavernier,J., Devos,R., Cornelis,S., Tuypens,T., der Heyden,J.V., Fiers,W. and Plaetinck,G. (1991) A human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific α chain and a β chain shared with the receptor for GM-CSF. *Cell*, 66, 1175–1184.

Torigoe,T., O'Connor,R., Santoli,D. and Reed,J.C. (1992) Interleukin-3 regulates the activity of the LYN protein-tyrosine kinase in myeloid-committed leukemic cell lines. *Blood*, 80, 617–624.

Ullrich,A. and Schlessinger,J. (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell*, 61, 203–212.

Vigon,I., Mornon,J.-P., Cocault,L., Mitjavila,M.-T., Tambourin,P., Gisselbrecht,S. and Souyri,M. (1992) Molecular cloning and characterization of *MPL*, the human homolog of the v-*mpl* oncogene: identification of a member of the hematopoietic growth factor receptor superfamily. *Proc. Natl Acad. Sci. USA*, 89, 5640–5644.

Ward,L.D., Howlett,G.J., Discolo,G., Yasukawa,Y., Hammacher,A., Moritz,R.L. and Simpson,R.J. (1994) The high affinity interleukin-6 (IL-6) receptor is a hexameric complex consisting of two molecules each of IL-6, IL-6 receptor and gp130. *J. Biol. Chem.*, 269, 23286–23289.

Watowich,S.S., Hilton,D.J. and Lodish,H. (1994) Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. *Mol. Cell. Biol.*, 14, 3535–3549.

Weiner,D.B., Liu,J., Cohen,J.A., Williams,W.V. and Greene,M.I. (1989) A point mutation in the *neu* oncogene mimics ligand induction of receptor aggregation. *Nature*, 339, 230–231.

Wheeler,E.F., Askew,D., May,S., Ihle,J.N. and Sherr,C.J. (1987) The v-*fms* oncogene induces factor-independent growth and transformation of the interleukin-3-dependent myeloid cell line FDC-P1. *Mol. Cell. Biol.*, 7, 1673–1680.

Yoshimura,A., Longmore,G. and Lodish,H.F. (1990) Point mutation in the exoplasmic domain of the erythropoietin receptor resulting in hormone-independent activation and tumorigenicity. *Nature*, 348, 647–649.

Young,D.C. and Griffin,J.D. (1986) Autocrine secretion of GM-CSF in acute myeloblastic leukemia. *Blood*, 68, 1178–1181.

Ziegler,S.F., Bird,T.A., Morella,K.K., Mosley,B., Gearing,D.P. and Baumann,H. (1993) Distinct regions of the human granulocyte-colony-stimulating factor receptor cytoplasmic domain are required for proliferation and gene induction. *Mol. Cell. Biol.*, 13, 2384–2390.

Received on January 24, 1995; revised on May 24, 1995

Note added in proof

We note a recent publication describing the activation of the mouse βc subunit by an amino-terminal truncation [J.Hanneman, T.Hara, M.Kawai, A.Miyajima, W.Osterlag and C.Stocking (1995) Sequential mutations in the interleukin-3 (IL-3)/granulocyte-macrophage-colony-stimulating factor/IL-5 receptor β-subunit genes are necessary for the complete conversion to growth autonomy mediated by a truncated βc subunit. *Mol. Cell. Biol.*, 15, 2402–2412].

RAPID COMMUNICATION

A Mutation of the Common Receptor Subunit for Interleukin-3 (IL-3), Granulocyte-Macrophage Colony-Stimulating Factor, and IL-5 That Leads to Ligand Independence and Tumorigenicity

By Richard D'Andrea, John Rayner, Paul Moretti, Angel Lopez, Gregory J. Goodall, Thomas J. Gonda, and Mathew Vadas

The cytokines interleukin-3, interleukin-5, and granulocyte-macrophage colony-stimulating factor bind with high affinity to a receptor complex that contains a ligand-specific α-chain and a common β-chain, hβc. We report here the isolation of a mutant form of hβc, from growth factor-independent cells, that arose spontaneously after infection of a murine factor-dependent hematopoietic cell line (FDC-P1) with a retroviral hβc expression construct. Analysis of this hβc mutation shows that a small (37 amino acid) duplication of extracellular sequence that includes two conserved sequence motifs is sufficient to confer ligand-independent growth on these cells and lead to tumourigenicity. Because this is a conserved region in the cytokine receptor superfamily, our results suggest that the large family of cytokine receptors has the capacity to become oncogenically active.

THE RECEPTOR SUBUNITS for interleukin-3 (IL-3), IL-5, and granulocyte-macrophage colony-stimulating factor (GM-CSF) are members of the cytokine receptor superfamily[1-3] that is characterized by a 200 amino acid extracellular module with a characteristic two β-barrel structure.[4,5] Within this family, the only extensive sequence conservation exists in a short membrane proximal extracellular sequence, Trp-Ser-Xaa-Trp-Ser (WSXWS), the role of which is unclear. Specific α-chains bind their cognate ligand with low affinity and combine with a shared β subunit, hβc, to form the high-affinity complex.[6-8] The stoichiometry of the active complex and the mechanisms mediating signalling have not been clearly established. hβc does not detectably bind cytokines by itself, but is crucial for signal transduction.[9] Two membrane proximal cytoplasmic sequences in the hβc cytoplasmic domain are important for proliferation in response to GM-CSF and induction of c-*myc*, whereas a more distal region of 140 amino acids has been shown to mediate activation of Ras, Raf1, MAP kinase, and p70 S6 kinase, and induction of c-*fos* and c-*jun*.[10,11] In addition, the cytoplasmic tyrosine kinase, JAK2, has recently been implicated in the IL-3 response and is rapidly phosphorylated and activated in response to ligand.[12] We have characterized a mutant, activated form of hβc that implicates a membrane proximal region of the receptor, including the WSXWS motif, as an important determinant in signalling. These results raise the possibility that other members of the cytokine receptor superfamily have oncogenic potential. Possible mechanisms involved in receptor activation are discussed.

MATERIALS AND METHODS

Construction of retrovirus expressing hβc. The construction of a full-length hβc cDNA has been described elsewhere.[13] The full-length coding sequence was excised from a pcDNA1hβc subclone and ligated into pRUF$_{NL}$Neo.[14]

Generation of factor-independent cells. Retroviral DNA was used to transfect an ecotropic packaging cell line, Ψ$_2$,[15] and virus from G418-resistant cells was used to infect the murine myeloid cell line, FDC-P1,[16] by cocultivation.[17] Cells were selected in G418 and, after several washes, incubated in medium with or without growth factor. After an extended time in culture (1 to 2 weeks), FDC-P1 cells from a wild-type hβc infection were clearly growing in the absence of any added growth factor. Parallel cocultivations with Ψ$_2$ cells producing hIL-3Rα or pRUF$_{NL}$Neo retroviruses showed no growth on FDC-P1 cells in the absence of factor. In subsequent experiments, we have not generated factor-independent cells from similar cocultivations with hβc retrovirus, implying that this mutant was generated spontaneously from a rare rearrangement during infection.

Isolation of genomic DNA and Southern analysis. Genomic DNA was isolated from cells using a Proteinase K/SDS procedure[18] and analyzed using a standard Southern protocol.[19] To detect hβc sequences, a ^{32}P-labeled probe was prepared by random primed oligo-extension (Multiprime Labelling System; Amersham, Arlington Heights, IL) of the hβc coding fragment excised and purified from a pcDNA1 subclone.

Polymerase chain reaction (PCR) and nucleotide sequencing. PCR was performed on 100 ng of genomic DNA using standard protocols.[20] The primers used for amplification were 5' TGGATCCTCTGTGGGTAGATCTGAGGCAG 3' and 5' TGAATTCCATAAAGAGCTCAGTGAACATCC 3'. Reactions were performed in a Perkin Elmer Thermocycler (Perkin Elmer Cetus, Norwalk, CT) and the cycling parameters were 30 cycles of 94°C for 1 minute, 57°C for 2 minutes, and 72°C for 3 minutes, with a final 10-minute extension in cycle 30. Reactions were denatured at 95°C for 5 minutes before cycling. PCR products were cloned into pGEM2 (Promega, Madison, WI) for DNA sequencing. Sequencing was performed using T7 Polymerase (Sequenase; US Biochemical Corp [USB], Cleveland, OH) as per the manufacturer's protocols.

Proliferation assays. FDC-P1 cells were infected as described and G418-resistant cells were maintained in the presence of growth factor (80 U/mL of murine GM-CSF). To assay proliferation, cells were washed and divided equally into medium with or without growth factor for 12 hours. Samples containing equivalent numbers

From the Hanson Centre for Cancer Research, Institute of Medical and Veterinary Science, Adelaide, Australia.

Submitted February 3, 1994; accepted February 25, 1994.

Supported in part by National Health & Medical Research Council Grant No. 920860 and National Institutes of Health Grant No. CA45822.

G.J.G., T.J.G., and M.V. contributed equally to this report.

Address reprint requests to Dr. Richard D'Andrea, Hanson Centre for Cancer Research, Institute of Medical and Veterinary Science, PO Box 14, Rundle Mall Post Office, Adelaide SA, 5000, Australia.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

0006-4971/94/8310-0044$3.00/0

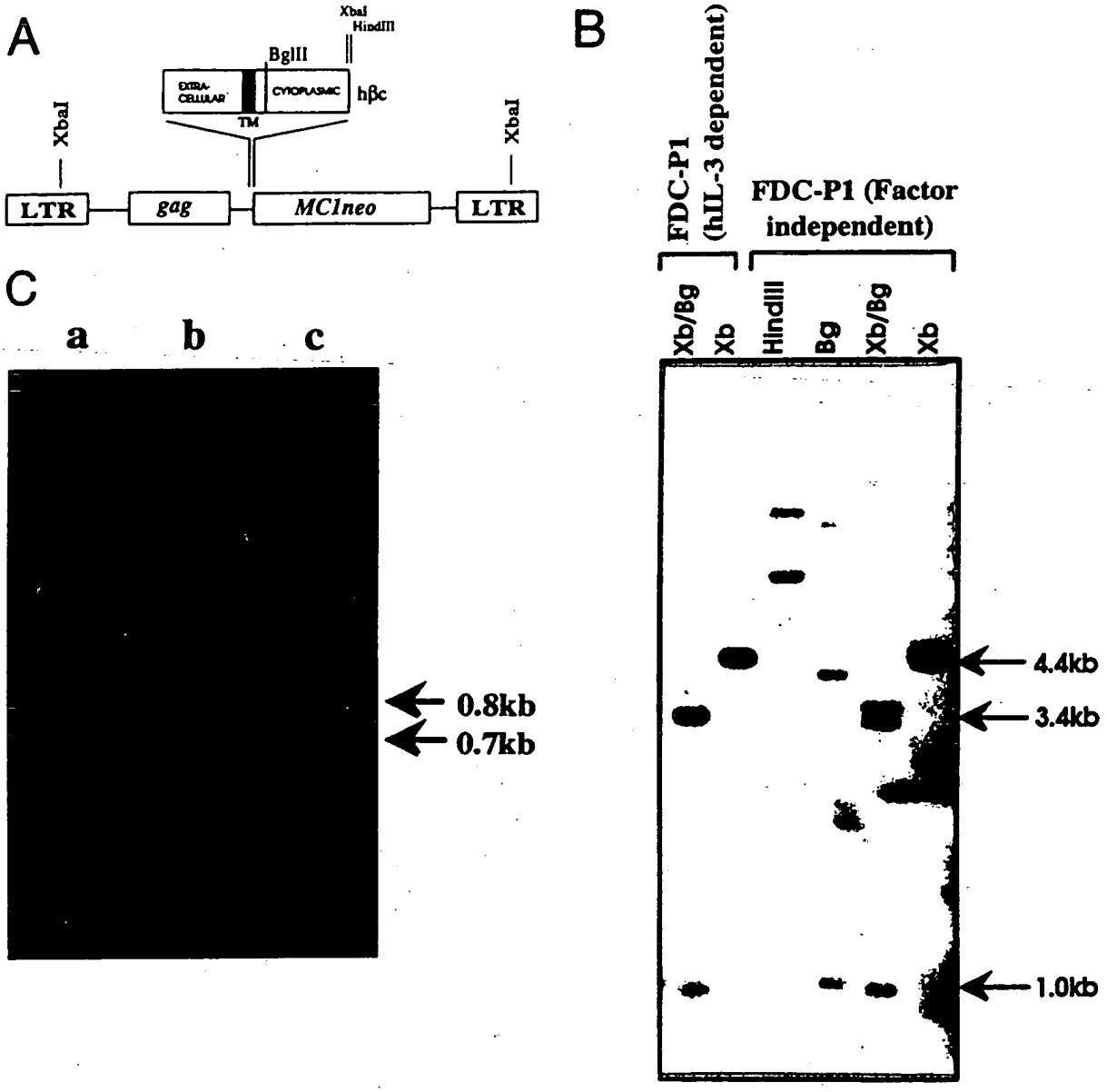


Fig 1. Generation and characterization of factor-independent cells. (A) Diagrammatic representation of the hβc retroviral construct. The vector used is pRUF$_{NL}$Neo.[14] Only relevant restriction sites are shown. (B) Southern analysis of human IL-3–dependent cells (FDC-P1 cells expressing hIL-3Rα and hβc) and the factor-independent hβc FDC-P1 cell line. Filters were probed with an hβc coding fragment. The wild-type hβc fragments are 4.4 kb in the *Xba* I digest and 3.4 kb and 1.0 kb in the *Xba* I/*Bgl* II digest. Clearly visible is a larger hβc fragment in the factor-independent FDC-P1 cells. (C) PCR from genomic DNA with internal hβc primers. The product generated spans the transmembrane domain and includes a small segment of cytoplasmic sequence. Lane a, markers, ϕX174 *Hae* III (Pharmacia, P-L Biochemicals Inc, Milwaukee, WI). Lane b, PCR using genomic DNA from factor-independent FDC-P1 cells as template. Lane c, PCR using genomic DNA from human IL-3–dependent FDC-P1 cells (as in [B]) as template. PCR products spanning the N-terminal and C-terminal regions of hβc were identical in both samples (data not shown).

of cells (4×10^5) were incubated in the presence of [^{3}H] thymidine and DNA synthesis measured using a standard protocol described previously.[21]

In vitro mutagenesis. For site-directed mutagenesis, the fragment to be mutagenized was cloned into pALTER (Promega) and single-stranded DNA prepared after infection of transformed *Escherichia coli* with helper phage (R408). DNA was prepared and mutagenesis performed with a mutagenic oligonucleotide following the protocols provided by the manufacturer.

RESULTS

Isolation and characterization of factor-independent cells. In the course f ur studies using FDC-P1 cells infected

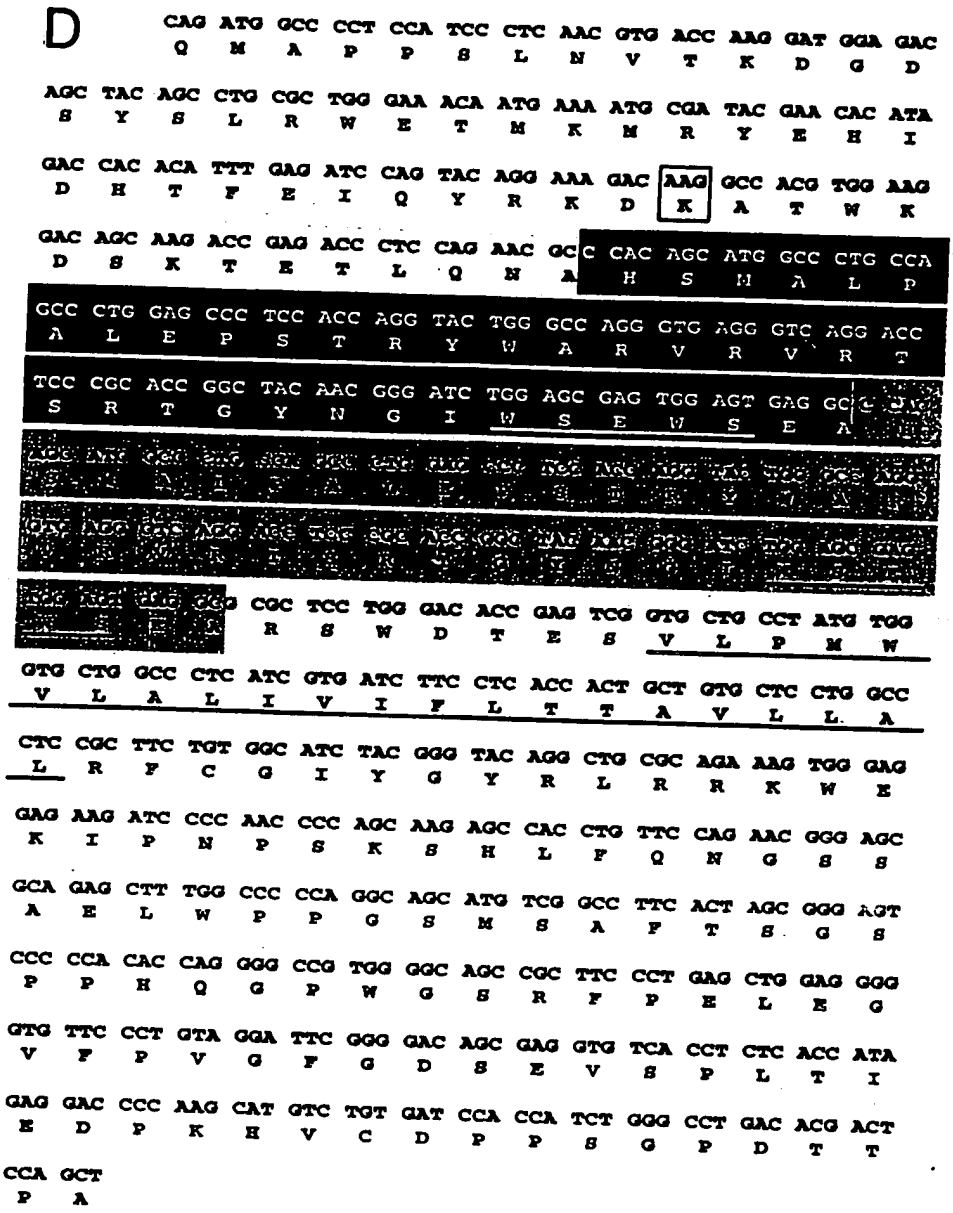


Fig 1. (Cont'd) (D) Sequence of the 0.8-kb hβc PCR product from factor-independent FDC-P1 cells showing the 111-bp tandem duplication in the extracellular domain. Only the sequence between the primers is shown. The transmembrane domain is indicated (underlined). The duplication leads to an in frame insertion resulting in a 37 amino acid duplicated segment (shaded) in the membrane proximal region of the molecule. This PCR product also contains a single amino acid substitution (T → K) (boxed). Two copies of the conserved WSXWS motif are also underlined.

with an hβc retrovirus, a cell line arose spontaneously that grew in the absence of mouse GM-CSF and IL-3 (see Materials and Methods and Fig 1A). To determine possible causes for this, we first examined cell supernatants for the presence of factors that could support the growth of FDC-P1 cells. Conditioned medium derived from the factor-independent cells did not support growth of untransfected cells, indicating that factor independence was not caused by autocrine production of growth factor. Secondly, we tested these cells for overexpression of hβc. However, immunostaining with the hβc-specific monoclonal antibody, 5F9, indicated that these cells were expressing similar levels of hβc to control human IL-3–dependent cells (data not shown). We characterized these cells further to determine whether there had been any alteration in the hβc cDNA on infection and generation of factor independence. Southern analysis indicated that two retroviral integrants were present in a factor-independent cell line (Fig 1B, *Hin*dIII digest). Digestion with a restriction enzyme that excises the proviral DNA (Fig 1B, *Xba* I and *Xba I/Bgl* II digests) indicated the presence f a larger than

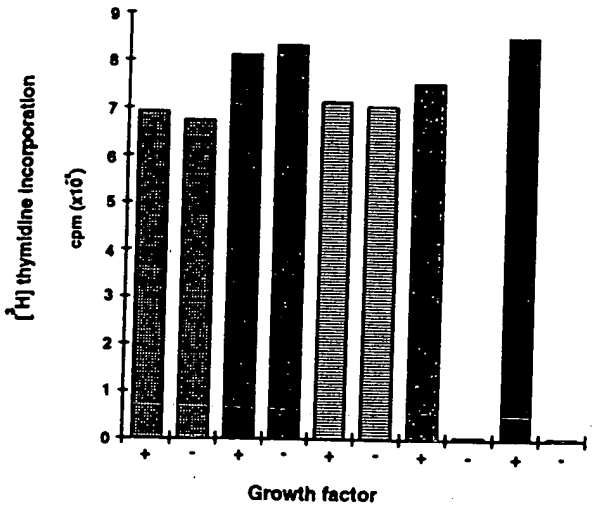


Fig 2. Proliferation of retrovirally infected FDC-P1 cells measured by [³H] thymidine incorporation.[21] (▨) Cells expressing hβc containing the duplicated segment and with the single amino acid alteration restored to wild-type; (▩) original factor-independent cells; (▩) cells expressing hβc from a retrovirus containing the recovered PCR fragment (see Fig 1C); (▩) cells infected with the pRUF$_{NL}$Neo vector; (▩) cells infected with a wild-type hβc retrovirus (Fig 1A). Presence or absence of growth factor (GM-CSF) is indicated.

normal internal fragment in one copy of the provirus. PCR amplification of DNA derived from the factor-independent cells using internal hβc primers indicated that there was an alteration giving rise to a larger internal hβc PCR product (Fig 1C, lane b), consistent with the Southern analysis. The altered PCR product spanned the transmembrane domain and membrane proximal segments of the receptor. Nucleotide sequencing of cloned PCR products generated from these cells indicated that the larger hβc product contained both a single point mutation (resulting in alteration of threonine to lysine) and a tandem duplication of 111bp; this resulted in a 37 amino acid duplicated segment in the extracellular portion of the receptor, proximal to the transmembrane domain (Fig 1D). We have reconstructed an hβc retrovirus containing this altered fragment. A full-length hβc receptor sequence was reconstructed in the retroviral vector, pRUF$_{NL}$Neo, using restriction fragments from hβc subclones and the 0.8-kb PCR fragment (see Fig 1C). Ψ_2 cells were transfected with the resultant plasmid, pRUF$_{NL}$βcFIΔ, and the G418-resistant producer cells used to introduce retrovirus into FDC-P1 cells. Cells infected with this retrovirus were capable of growth in the absence of added growth factor. Cells infected with control virus (either vector alone r hβc retrovirus) gave no growth in the absence of added growth factor (Fig 2). To show that the duplication alone could c nfer factor-independent growth, we cloned the reconstructed hβc into pALTER (Promega) and restored the altered lysine residue to threonine by site-directed mutagenesis. FDC-P1 cells selected f r G418 resistance after infecti n with this retrovirus (pRUF$_{NL}$βcFI$_{T380}$) were again factor-independent, indicating that the point mutation was n t involved in generating factor independence in these cells. Cells infected with mutant and control retrovirus were also plated in agar with and without growth factor. Only cells infected with pRUF$_{NL}$βcFIΔ formed G418-resistant colonies in agar in the absence of added factor (data not shown), consistent with expression of the mutant hβc being sufficient to confer factor-independent growth on FDC-P1 cells. Clearly, factor independence in the original cell line is not conferred as a result of retroviral integration or sequence alteration in another region f hβc. We can also rule out a requirement for the normal hβc product, derived from the second retroviral insertion, in the factor-independent phenotype of the original cell line. We interpret these results as indicating that the alteration in receptor sequence has led to constitutive activation of the hβc subunit.

Tumorigenicity. Although members of the cytokine receptor superfamily have not yet been associated clinically with oncogenesis, an activated form of the murine erythropoietin receptor (mEpoR) has been reported and shown to have tumourigenic potential.[22,23] We therefore tested the tumourigenic potential of the mutant hβc molecule by injecting factor-independent and control FDC-P1 cells into syngeneic mice (Table 1). Control cells, infected with the pRUF$_{NL}$Neo vector or with a wild-type hβc retrovirus, were not tumourigenic. However, FDC-P1 cells, infected with pRUFβcFIΔ (a retrovirus expressing the mutant form of hβc), generated solid tumors with a short latency in all mice injected. This result shows the tumourigenic capacity of activated hβc and emphasises the oncogenic potential of the large family of cytokine receptors.

DISCUSSION

We have shown that a small (37 amino acid) duplication of extracellular hβc sequence is sufficient to confer ligand-independent growth on FDC-P1 cells and leads to tumourigenicity. We believe that the sequence defined by this duplication may be important in signalling because the altered receptor structure must in some way mimic a ligand-induced signalling event. Structural predictions suggest that the seg-

Table 1. Tumourigenicity of Transfected FDC-P1 Cells

Cells Injected	Tumors Observed/No. of Mice Injected
HBBS	0/3
FDC-P1 (pRUFNeo)	0/4
FDC-P1 (pRUFhβc)	0/4
FDC-P1 (pRUFhβcFIΔ)	5/5

FDC-P1 cells were infected with wild-type and mutant hβc retrovirus (indicated in parentheses) and selected for G418 resistance. To assess tumourigenicity, cells were injected into syngeneic DBA2 male mice (8 weeks old). Cells (5×10^6 per mouse) were washed in serum-free medium and resuspended in Hanks' Balanced Salt Solution (HBBS; 1×10^7 cells/mL) and 0.5 mL was injected into mice subcutaneously at a single site. Mice were observed for 2 to 3 months for signs of palpable or visible tumours at the site of injection. Tumourigenic cells gave rise to visible masses with a short latency after injection (2 to 3 weeks). Mice with solid tumors were killed and autopsies were performed to assess the extent of tumor formation. Representative mice injected with nontumorigenic cells were also killed and post mortems were performed to ensure that tumors had not developed.

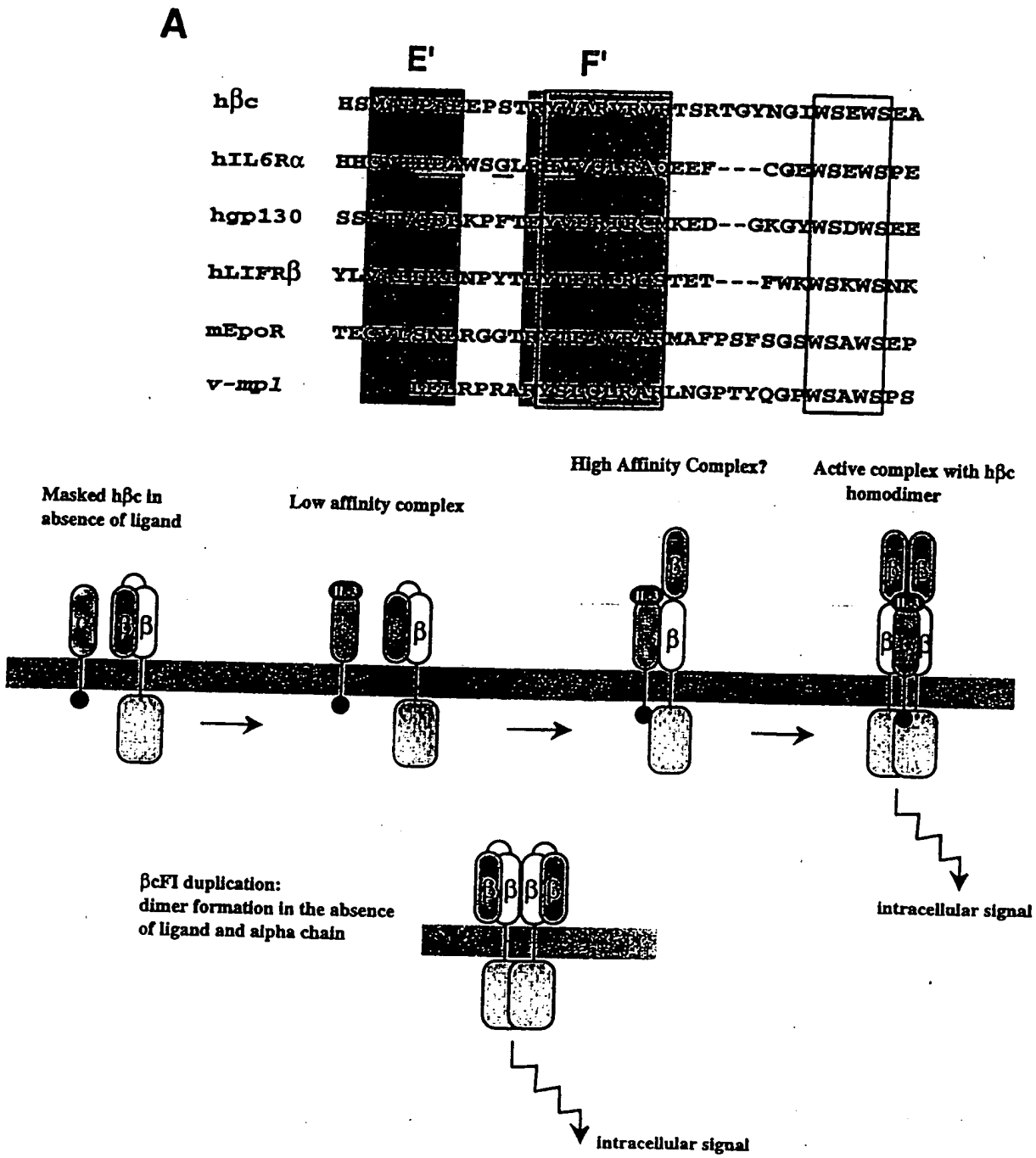


Fig 3. Possible role of the duplicated segment in hβc activation. (A) Alignment of receptor amino acid sequences with the duplicated segment from the mutant hβc. The predicted location of β-strands E' and F' (see Bazan[4]) is shown with light grey shading. Amino acids shown to be important in association of IL-6Rα with gp130[31] are underlined. Conserved sequence motifs discussed in the text are boxed. Only the cytokine receptor segment of v-*mpl* is shown. Sequences have already been described.[6,33,35,37,38,39] (B) Model for formation of a functional hβc receptor complex. Constitutive activation may involve the formation of an exposed site capable of mediating dimerization through association with an equivalent region of a second similar hβc molecule. We speculate that, in the absence of ligand, the membrane proximal region of hβc is masked and hence prevented from association and dimer formation. In the presence of ligand, an intermediate complex may be the α-chain hβc heterodimer. Association with the α-chain/ligand complex may relieve masking and allow hβc dimerization. Whether α-chain remains associated with the active complex is not known. In mutant hβc, the duplicated segment may permit the membrane proximal domains of mutant hβc monomers to interact in the absence of ligand or α-chain.

ment duplicated comprises two β-strands (E′ and F′) and two connecting loops, E′-F′ and F′-G′.[4,24] It also includes the WSXWS motif and an adjacent conserved basic region (consensus YXXXVRXR, see Patthy[25]) corresponding to the predicted β-strand F′ (see Fig 3A). The following mechanisms of activation are possible: (1) activation occurring through association of mutant hβc with murine receptor α-chains expressed on these cells; (2) activation through an altered tertiary structure that mimics a ligand-induced event; and (3) the duplicated segment allows hβc dimer formation in the absence of ligand and hence plays a direct role in activation of the mutant receptor. The duplicated segment may form an exposed site mediating interaction of two hβc molecules. In the normal receptor, the single copy of this site maybe involved in dimer formation, but only in the presence of specific α-chain and bound ligand (Fig 3B). Indirect support for this third model ([3] above) comes from recent studies with gp130, the signalling subunit for IL-6, CNTF, and LIF. Based on the fact that receptor complexes for IL-3, IL-5, and GM-CSF share a signalling subunit, as do those for IL-6, Oncostatin M, CNTF, and LIF, it has been suggested that hβc may form a dimer in a similar fashion to gp130.[26,27] The observation that the signalling subunits, hβc and gp130, have homology in their intracellular domains[10] further suggests that there are similarities in the mechanisms mediating signalling. Additional support for this model comes from a study with a chimeric receptor containing the extracellular domain of the mEpoR and the intracellular domain of a murine hβc homologue, AIC2A.[28] This chimera mediates Epo-dependent proliferation consistent with the activated form of AIC2A being a dimer.

Several studies on other receptors within this superfamily suggest the importance of this conserved membrane proximal segment in signal transduction. The role of the WSXWS motif has been extensively studied, although results are conflicting. Whereas mutations in this sequence have been reported that abolish ligand binding,[29] mutations in the WSXWS motif of mEpoR suggest a role in signalling.[30] A comprehensive mutagenesis of the IL-6 receptor α-chain (IL-6Rα) identified seven mutations that appear to specifically abolish signalling by preventing association with a signalling subunit, gp130.[31] All except one of these lie in strands E′ and F′ and the adjoining loop (underlined in Fig 3A), implying that this region is important in receptor subunit association. These observations, together with the recent evidence that IL-6Rα complexes with a gp130 dimer,[26,32] suggest that there could be a region in IL-6Rα that includes the conserved β-strand F′ (and perhaps the WSXWS motif) that is important in interaction with one of the gp130 molecules. Perhaps in hβc this region is important in dimer formation.

The hβc receptor contains two of the characteristic cytokine receptor modules[8] and we speculate that the N-terminal receptor structure masks a membrane proximal site involved in dimerization, thereby preventing association. Other cloned cyt kine receptor subunits, LIFRβ[33] and *MPL*,[34] resemble the hβc structure in that they have a repeated cytokine receptor module. Truncation from the N-terminus may be expected t activate these receptors. Although we have not yet tested N-terminal truncations of hβc, the onc gene v-*mpl*[35] represents a truncation of this type. It is notable that the extracellular region remaining in the v-*mpl* product includes the conserved segment duplicated in the hβc mutant (see Fig 3A), consistent with some of these sequences being essential for signalling and perhaps contributing to formation of a dimeric complex.

In conclusion, the activated form of hβc suggests that a membrane proximal extracellular region of cytokine receptor subunits has an important role in signalling and we suggest that this region may be critical for dimerization (Fig 3B). Duplication of this segment may allow hβc association in the absence of ligand and hence constitutively activate the receptor. Indeed, mEpoR is activated by forced dimerization through a mutation that introduces a cysteine residue.[36] This constitutive EpoR mutation and the hβc mutant described here suggest that mutations leading to inappropriate receptor subunit association could be a possible mode of oncogenic activation for several cytokine receptor family members.

ACKNOWLEDGMENT

We thank Chris Bagley for his comments on receptor structure and his critical reading of the manuscript. We also thank Liz Macmillan for technical advice throughout the course of this work and Bronny Cambareri for the proliferation assays. Brendan Jenkins assisted with site-directed mutagenesis.

REFERENCES

1. Miyajima A, Kitamura T, Harada N, Yokota T, Arai K: Cytokine receptors and signal transduction. Annu Rev Immunol 10:295, 1992

2. Cosman D, Lyman SD, Idzerda RL, Beckmann MP, Park LS, Goodwin RG, March CJ: A new cytokine receptor superfamily. Trends Biochem Sci 15:265, 1990

3. Miyajima A, Hara T, Kitamura T: Common subunits of cytokine receptors and the functional redundancy of cytokines. Trends Biochem Sci 17:378, 1992

4. Bazan JF: Structural design and molecular evolution of a cytokine receptor superfamily. Proc Natl Acad Sci USA 87:6934, 1990

5. De Vos AM, Ultsch M, Kossiakoff AA: Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. Science 255:306, 1992

6. Kitamura T, Sato N, Arai K, Miyajima A: Expression cloning of the human IL-3 receptor cDNA reveals a shared β subunit for the human IL-3 and GM-CSF receptors. Cell 66:1165, 1991

7. Tavernier J, Devos R, Cornelis S, Tuypens T, Van der Heyden J, Fiers W, Plaetinck G: A human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific chain and a β chain shared with the receptor from GM-CSF. Cell 66:1175, 1991

8. Hayashida K, Kitamura T, Gorman DM, Arai K, Yokota T, Miyajima A: Molecular cloning of a second subunit of the receptor for human granulocyte macrophage colony-stimulating factor (GM-CSF): Reconstitution of a high affinity receptor. Proc Natl Acad Sci USA 87:9655, 1990

9. Kitamura T, Miyajima A: Functional reconstitution of the human interleukin-3 receptor. Blood 80:84, 1992

10. Sakamaki K, Miyajima I, Kitamura T, Miyajima A: Critical cytoplasmic domains of the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors for growth signal transduction and tyrosine phosphorylation. EMBO J 11:3541, 1992

11. Sato N, Sakamaki K, Terada N, Arai K-I, Miyajima A: Signal transduction by the high-affinity GM-CSF receptor: Two distinct cytoplasmic regions of the common β subunit responsible for different signaling. EMBO J 12:4181, 1993

12. Silvennoinen O, Witthuhn A, Quelle FW, Cleveland JL, Yi T, Ihle JN: Structure of the murine JAK2 protein tyrosine kinase and its role in IL-3 signal transduction. Proc Natl Acad Sci USA (in press)

13. Barry SC, Bagley CJ, Phillips J, Dottore M, Cambareri B, Moretti P, D'Andrea R, Goodall GJ, Shannon MF, Vadas MA, and Lopez AF: Two contiguous residues in human interleukin-3, Asp^{21} and Glu^{22}, selectively interact with the α- and β-chains of its receptor and participate in function. J Biol Chem 269:8488, 1994

14. Rayner J, Gonda TJ: A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. Mol Cell Biol 14:880, 1994

15. Mann R, Mulligan RC, Baltimore D: Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33:153, 1983

16. Dexter TM, Garland J, Scott D, Scolnick E, Metcalf D: Growth of factor-dependent hemopoietic precursor cell lines. J Exp Med 152:1036, 1980

17. Lang RA, Metcalf D, Gough NM, Dunn AR, Gonda TJ: Expression of a hemopoietic growth factor cDNA in a factor-dependent cell line results in autonomous growth and tumorigenicity. Cell 43:531, 1985

18. Hughes SH, Payvar F, Spector D, Schimke RT, Robinson HL, Payne GS, Bishop JM, Varmus HE: Heterogeneity of genetic loci in chickens: Analysis of endogenous viral and normal genes by cleavage of DNA with restriction endonucleases. Cell 18:347, 1979

19. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory, 1989

20. Saiki RK: Amplification of genomic DNA, in Innes MA, Gelfand DH, Sninsky JJ, White TJ (eds): PCR Protocols: A Guide to Methods and Applications. San Diego, CA, Academic, 1990, p 13

21. Lopez AF, Dyson PG, To LB, Elliot MJ, Milton SE, Russell JA, Juttner CA, Yang YC, Clark SC, Vadas MA: Recombinant human IL-3 stimulation of hematopoiesis in humans: Loss of responsiveness with differentiation in the neutrophilic myeloid series. Blood 72:1797, 1988

22. Yoshimura A, Longmore G, Lodish HF: Point mutation in the exocytoplasmic domain of the erythropoietin receptor resulting in hormone-independent activation and tumorigenicity. Nature 348:647, 1990

23. Longmore GD, Lodish HF: An activating mutation in the murine erythropoietin receptor induces erythroleukemia in mice: A cytokine receptor superfamily oncogene. Cell 67:1089, 1991

24. Goodall GJ, Bagley CJ, Vadas MA, Lopez AF: A model for the interaction of the GM-CSF, IL-3 and IL-5 receptors with their ligands. Growth Factors 8:87, 1993

25. Patthy L: Homology of a domain of the growth hormone/prolactin receptor family with type III modules of fibronectin. Cell 61:13, 1990 (letter)

26. Davis S, Aldrich TH, Stahl N, Pan L, Taga T, Kishimoto T, Ip NY, Yancopoulos GD: LIFRβ and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. Science 260:1805, 1993

27. Stahl N, Yancopoulos G: The alphas, betas and kinases of cytokine receptor complexes. Cell 74:587, 1993

28. Chiba T, Nagata Y, Machide M, Kishi A, Amanuma H, Sugiyama M, Todokoro K: Tyrosine kinase activation through the extracellular domains of cytokine receptors. Nature 362:646, 1993

29. Miyazaki T, Maruyama M, Yamada G, Hatakeyama M, Taniguchi T: The integrity of the conserved "WS motif" common to IL-2 and other cytokine receptors is essential for ligand binding and signal transduction. EMBO J 10:3191, 1991

30. Quelle DE, Quelle FW, Wojchowski DM: Mutations in the WSAWSE and cytosolic domains of the erythropoietin receptor affect signal transduction and ligand binding and internalisation. Mol Cell Biol 12:4553, 1992

31. Yawata H, Yasukawa K, Natsuka S, Murakami M, Yamasaki K, Hibi M, Taga T, Kishimoto T: Structure-function analysis of human IL-6 receptor: Dissociation of amino acid residues required for IL-6 binding and for IL-6 signal transduction through gp130. EMBO J 12:1705, 1993

32. Murakima M, Hibi M, Nakagawa N, Nakagawa T, Yasukawa K, Yamanishi K, Taga T, Kishimoto T: IL-6 induced homodimerization of gp130 and associated activation of a tyrosine kinase. Science 260:1808, 1993

33. Gearing DP, Thut CJ, VandenBos T, Gimpel SD, Delaney PB, King J, Price V, Cosman D, Beckmann MP: Leukemia inhibitory factor receptor is structurally related to the IL-6 signal transducer, gp130. EMBO J 10:2839, 1991

34. Vigon I, Mornon J-P, Cocault L, Mitjavila M-T, Tambourin P, Gisselbrecht S, Souyri M: Molecular cloning and characterisation of *MPL*, the human homolog of the v-*mpl* oncogene: Identification of a member of the haematopoietic growth factor receptor superfamily. Proc Natl Acad Sci USA 89:5640, 1992

35. Souyri M, Vigon I, Penciolelli J-F, Heard J-M, Tambourin P, Wendling F: A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalises hematopoietic progenitors. Cell 63:1137, 1990

36. Watowich SS, Yoshimura A, Longmore GD, Hilton DG, Yoshimura Y, Lodish HF: Homodimerization and constitutive activation of the erythropoietin receptor. Proc Natl Acad Sci USA 89:2140, 1992

37. Yamasaki T, Taga T, Hirata Y, Yawata H, Kawanishi Y, Seed B, Tadatsugu T, Hirano T, Kishimoto T: Cloning and expression of the human interleukin-6 (BSF-2/IFNβ 2) receptor. Science 241:825, 1988

38. Hibi M, Murakami M, Saito M, Hirano T, Taga T, Kishimoto T: Molecular cloning and expression of an IL-6 signal transducer, gp130. Cell 63:1149, 1990

39. D'Andrea AD, Lodish HF, Wong GG: Expression cloning of the murine erythropoietin receptor. Cell 57:277, 1989

RAPID COMMUNICATION

Extracellular Truncations of hβc, the Common Signaling Subunit for Interleukin-3 (IL-3), Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF), and IL-5, Lead to Ligand-Independent Activation

By Richard J. D'Andrea, Simon C. Barry, Paul A.B. Moretti, Karen Jones, Sarah Ellis, Mathew A. Vadas, and Gregory J. Goodall

The hypothesis that extracellular truncation of the common receptor subunit for interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor, and IL-5 (hβc) can lead to ligand-independent activation was tested by infecting factor-dependent hematopoietic cell lines with retroviruses encoding truncated forms of hβc. A truncation, resembling that in v-Mpl, and retaining 45 hβc-derived extracellular residues, led to constitutive activation in the murine myeloid cell line, FDC-P1. However, infection of cells with retrovirus encoding a more severely truncated receptor, retaining only 7 hβc-derived extracellular residues, did not confer factor independence on these cells. These experiments show that truncation activates the receptor and define a 37-amino acid segment of hβc (H395-A431) which contains two motifs conserved throughout the cytokine receptor superfamily (consensus Y/H XX R/Q VR and WSXWS), as essential for factor-independent signaling. The mechanism of activation was also investigated in less severe truncations. A receptor that retains the entire membrane-proximal domain (domain 4) also conferred factor independent growth on FDC-P1 cells; however, a retrovirus encoding a truncated form of hβc having two intact membrane proximal domains did not have this ability, suggesting that domain 3 may have an inhibitory role in hβc. The ability of these receptors to confer factor independence was cell specific as demonstrated by their inability to confer factor-independent growth when introduced into the murine IL-3–dependent pro-B cell line BaF-B03. These results are consistent with a model in which activation requires unmasking of an interactive receptor surface in domain 4 and association with a myeloid-specific receptor or accessory component. We suggest that in the absence of ligand intramolecular interactions prevent inappropriate signaling.

RECEPTORS for hematopoietic growth factors (HGFs) are members of the expanding cytokine receptor superfamily that is characterized by a 200-amino acid extracellular receptor module (CRM) composed of two discrete folding domains, each of which contains seven β strands folded into antiparallel β sandwiches, and bears a structural similarity to the fibronectin type III module and the Ig constant domains.[1,2] This structure has been confirmed for two members of the family, the growth hormone and prolactin receptors, by x-ray crystallographic studies.[3,4] In addition to this similarity in tertiary structure, members of this superfamily share a number of conserved sequence elements: (1) four conserved cysteine residues located in the N-terminal domain, (2) a membrane proximal Trp-Ser-Xaa-Trp-Ser (where Xaa is any amino acid) motif, also known as the "WSXWS box," located in the C-terminal domain,[1,5,6] and (3) a proline-rich motif (PRM), which may be involved in signal transduction, located in the membrane proximal region of the cytoplasmic domain of most receptors in this family.[7,8] Members f this family of receptors do not contain recognizable tyrosine kinase domains; signaling depends on association with cytoplasmic tyrosine kinases of the JAK family and the subsequent activation of signal transducers and activators of transcription.[9,10]

Constitutive mutations in cytokine receptors have provided insight into the process of receptor activation. Activating mutations may act by mimicking the structure of the normal receptor in the ligand-activated state and therefore could provide important clues to the activation process. Given the normal role of cytokines in cell proliferation and survival these activating mutations may be predicted to hav oncogenic potential. Indeed, the murine v-*mpl* oncogene encodes a constitutively activated cytokine receptor that has been transduced by the murine myeloproliferative leukemia virus (MPLV). v-Mpl is a 284-amino acid fusion protein in which the first 100 amino acids are derived from the region and the other 184 amino acids are from c-*mpl*,[11] which encodes the receptor for thrombopoietin (TPO).[12-16] The extracellular domain is mostly composed of the F-MuLV *env* sequences, but with 43 amino acids derived from c-Mpl (Fig 1A). Thus, in v-Mpl most of the TPO-R extracellular domain has been deleted, generating a truncated receptor. Although the mechanism of activation of the v-*mpl* product has not been determined it has been shown that Mpl can be activated by cysteine substitutions in the predicted dimer interface domain, suggesting one normal mechanism of activation probably involves ligand-induced homodimerization.[17]

Forced receptor homodimerization is involved in aberrant activation of other cytokine receptors. A constitutively active form of the erythropoietin receptor (EPO-R) containing an arginine-to-cysteine substitution at position 129 has been described. This R129C form of EPO-R forms disulfide-linked homodimers in the absence of EPO, suggesting that wild-type EPO-R is also activated by ligand-induced homodimerization.[18] Infection of mice with recombinant retrovirus

From the Division of Human Immunology, Hanson Centre for Cancer Research, Adelaide, Australia.

Submitted December 18, 1995; accepted December 26, 1995.

Supported by grants from the National Health and Medical Research Council of Australia and the Anti-Cancer Foundation of the Universities of South Australia. P.A.B.M. is a recipient of a Dora Lush Postgraduate Scholarship from the National Health and Medical Research Council of Australia.

Address reprint requests to Dr Richard D'Andrea, Division of Human Immunology, Hanson Centre for Cancer Research, Frome Rd, Adelaide, 5000 Australia.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

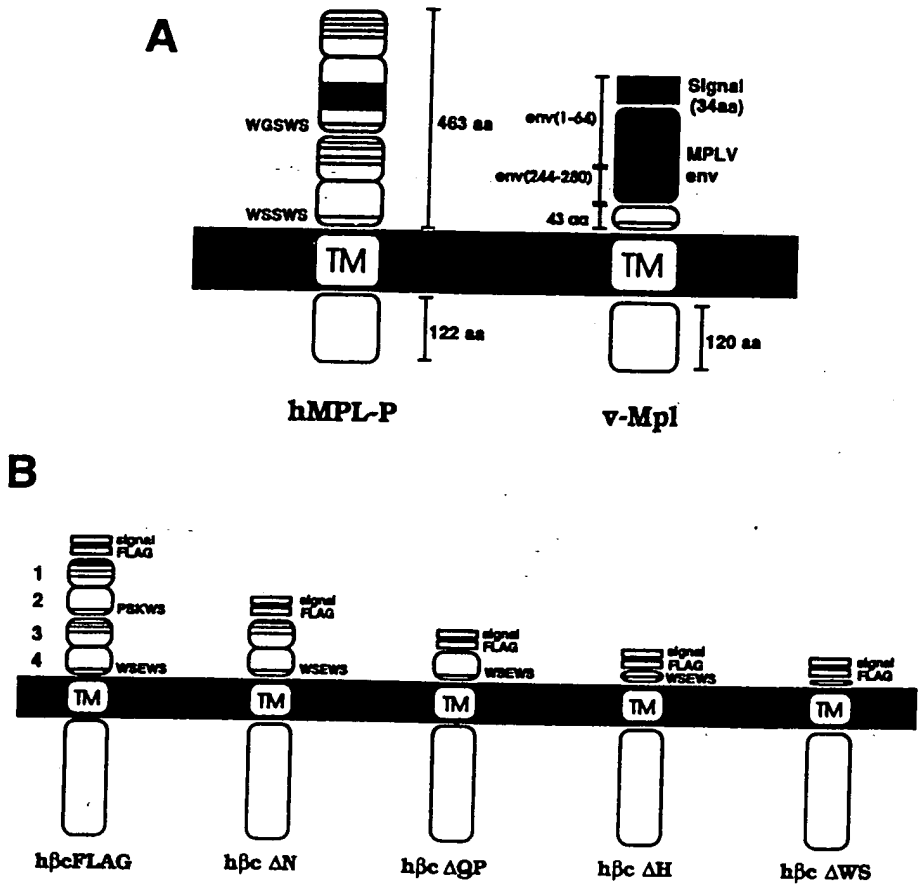


Fig 1. Cytokine receptor extracellular truncations. (A) Schematic illustration showing the major isoform of Mpl (hMpl-P) and the v-Mpl fusion product. Mpl has two CRMs, the distal one containing an insertion of 50 residues. The v-Mpl product is derived from the MPLV retrovirus and is a fusion protein containing rearranged *env* sequences fused to truncated Mpl. The conserved cysteine residues and WSXWS motif characteristic of the cytokine receptor family are indicated by horizontal lines. (B) Schematic illustration showing the extracellular truncation mutants of hβc that have been generated. Site-directed mutagenesis was used to construct in-frame deletions of the FLAG-hβc cDNA. Residues deleted are as follows: hβcΔN, E25-E232, hβcΔQP, E25-N337; hβcΔH, E25-A394; hβcΔWS, E25-A431. The signal sequence and position of the introduced FLAG octapeptide are indicated. Conserved cysteine residues and the WSXWS sequence that are characteristic of the cytokine receptor family are shown by horizontal lines. Extracellular domains are numbered from 1 to 4.

expressing this receptor induces erythroleukemia.[19] The introduction of other cysteine residues into the EPO-R membrane proximal domain also leads to disulfide-linked homodimers that are constitutively active.[20] A mutation in the transmembrane domain of the human common β subunit (hβc) for interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor (GM-CSF), and IL-5, which presumably works through a similar mechanism, has recently been described.[21] In this mutant (hβcV449E), glutamic acid is substituted for valine 449 in the transmembrane domain, conferring factor-independent growth on FDC-P1 and BaF-B03 cells. This mutation is analogous to an activating mutation in the c-*neu* proto-oncogene (*HER*-2, erb-B2 receptor), which leads to receptor oligomerization via a transmembrane-mediated association.[22,23] An activated form of the murine βc subunit has recently been isolated from a spontaneous, factor-independent subline of the promyelocytic cell line, D35. In this receptor the extracellular domain has been replaced by 34 amino acids encoded by intron 10 sequences and it is postulated that the replacement of the normal extracellular sequences with the intron encoded segment facilitates homodimerizati n, perhaps through an extracellular cysteine bridge.[24]

We have recently described two ligand-independent mutations in the membrane proximal, extracellular domain f hβc that suggest an alternative mechanism of aberrant cyt kine receptor activation. One f these (hβcFIΔ) is a duplication of a short (37-amino acid) receptor segment including the WSXWS motif[25] and the second (hβcI374N) is a point mutation leading to a single amino acid substitution (isoleucine 374 to asparagine).[21] Both mutant receptors confer growth factor independence and tumorigenicity on FDC-P1 cells. The mechanism of activation for these mutations is less clear; however, they do not confer ligand-independent proliferation in BaF-B03 cells, suggesting a different mechanism to hβcV449E.[21] We have proposed that these mutations may lead to an alteration in receptor structure that results in the unmasking of an interactive surface that is not normally available in the receptor monomer.[21,25] We speculated further that ligand association leads to activation by exposing such a surface and allowing an interaction that can lead to signaling. One prediction from such a model is that extracellular truncation may lead to activation by unmasking an interactive region and it is notable that the severely truncated v-Mpl product retains 43 amino acids of extracellular sequence,[11] including the two conserved motifs duplicated in hβcFIΔ (Fig 1A).

In this context we have now tested the hypothesis that extracellular truncati n could lead to constitutive activation of cytokine recept rs. We have examined a series of hβc extracellular recept r truncations for their capacity t confer growth in factor-dependent cell lines in the absence of growth factor. We find that infection of the murine myeloid cell line, FDC-P1, with retroviruses encoding truncated

forms of hβc leads to factor-independent proliferation provided domains 1-3 are removed and a conserved membrane proximal segment is retained.

MATERIALS AND METHODS

Construction and expression of hβc receptor mutations. The hβc cDNA forms used in this study were constructed by in vitro mutagenesis. We utilized a mutagenic oligonucleotide (5'CCCTGTGCTGGGTGCTGAGCGGCGCACAGGCAGACTACAAGGACGACGACGACAAGGAAGAAACCATCCCG3') to incorporate the coding sequence for the FLAG octapeptide (Eastman-Kodak Co, New Haven, CT), DYKDDDDK, after the signal sequence (between residues A24 and E25) in the hβc cDNA. To optimize cleavage of the signal peptide, residues 18-24 were altered from ERSLAGA to VLSGAQA. These modifications were confirmed by sequencing and this FLAG-hβc cDNA cloned into pALTER (Promega, Madison, WI) for further mutagenesis. Mutagenesis was performedwith mutagenic oligonucleotides designed to remove the desired extracellular sequence while leaving coding sequences for the N-terminal signal sequence and the inserted FLAG octapeptide intact. Mutagenic oligos were as follows: hβcΔN; 5'GACGACGACGACAAGGTTTGTTGGGACTCC3', hβcΔH; 5' GACGACGACGACAAGCACAGCATGGCCCTG3', hβcΔQP; 5'ACGACGACGACAAGATCCAGATGGCCCCT3', hβcΔWS; 5'ACGACGACGACAAGCGCTCCTGGGACACC3'. All mutants were obtained after screening DNA minipreparations from the ampicillin resistant colonies for an hβc fragment of diagnostic size. Mutations were confirmed by DNA sequencing, which was performed using T7 Polymerase (Sequenase; United States Biochemical Corp [USB, Cleveland, OH]) as per the manufacturer's protocols. The receptor mutants generated in this study are described in detail in the text and are shown schematically in Fig 1B. For expression, mutant receptors were excised from pALTER and cloned into the retroviral expression vector pRUFNeo.[26] Retroviral DNA was used to transfect the ecotropic packaging cell line, ψ2,[27] and virus from G418-resistant cells was used to infect the murine hematopoietic cell lines, FDC-P1[28] or BaF-B03[29] by cocultivation.[30]

Cell lines, growth factors, and flow cytometry. Murine myeloid, IL-3/GM-CSF–dependent FDC-P1 cells, and derived cell lines were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 7.5% fetal bovine serum and murine GM-CSF (80 U/mL; gift from Dr G. Begley, Walter and Eliza Hall Institute for Medical Research, Melbourne Australia). Murine IL-3–dependent BaF-B03 cells were maintained in DMEM supplemented with 7.5% fetal bovine serum and murine IL-3 (300 U/mL; gift from Dr A. Hapel, John Curtin School for Medical Research, Canberra, Australia). After cocultivation with producer cells transfected with pRUFNeo hβc constructs, FDC-P1, or BaF-B03 cells were harvested and selected in the presence of G418 (1 mg/mL; Sigma-Aldrich Pty Ltd, Castle Hill, NSW Australia) and growth factor. Receptor expression was examined by flow cytometric analysis after staining with the FLAG-specific monoclonal antibody (MoAb), M2 (Eastman-Kodak Co). Briefly, cells were washed and resuspended in cold phosphate-buffered saline supplemented with 5% bovine serum albumin (PBSA). Cells were incubated with the M2 MoAb (1:300) for 20 minutes on ice, washed, and subsequently incubated with biotinylated antimouse IgG (1:50; Vector Lab Inc, Burlingame, CA) for 20 minutes on ice. After washing and resuspension in cold PBSA the cells were incubated with streptavidin conjugated phycoerythrin (1:50; Caltag Laboratories, San Francisco, CA) for a further 20 minutes, washed, resuspended in PBSA + 0.01% sodium azide, and analyzed using an Epics-Profile II analyzer (Coulter, Hialeah, FL).

Mitogenic assays. To assay proliferation, infected cells (FDC-P1, BaF-B03) were washed three times with PBS to remove growth [illegible] cells per well] in a 96-well microtiter plate in the presence and absence of the appropriate growth factor. Cell growth was determined after 72 hours using a Cell titer 96 nonradioactive cell proliferation assay (Promega, Madison, WI). Quantitation was performed using an automated plate reader (BioRad Laboratories, Pty Ltd, North Ryde, NSW Australia). To establish relative growth rates of cells expressing factor-independent mutants, cells were first selected for factor independence, seeded at 5,000 cells per well, and proliferation assayed, as described earlier, on each of the following 4 days.

Preparation of cellular DNA and polymerase chain reaction (PCR) analysis. Genomic DNA was isolated from cells using a Proteinase K/SDS procedure.[31] PCR was performed on 300 ng of genomic DNA using standard protocols.[32] Internal hβc primers were used for amplification (5'TGAATTCGCCTGTCCAGAGCTGACCAGGG 3' and 5'TGGATCCTCTGTGGGTAGATCTGAGGCAG 3'). Reactions were performed in a Perkin Elmer Thermocycler and the cycling parameters were: 30 cycles—94°C for 1 minute, 60°C for 1 minute, 72°C for 1 minute. Reactions were denatured at 94°C for 2 minutes before cycling. PCR products were fractionated using 1% agarose mini-gels and visualized by staining with ethidium bromide.

RESULTS

Construction and expression of mutant hβc cDNA. In vitro mutagenesis was used to construct the series of extracellular hβc deletion mutants indicated in Fig 1B. In each of these mutations a large portion of extracellular region was replaced by the FLAG octapeptide. The N-terminal FLAG octapeptide does not affect receptor binding or function in FDC-P1 cells (data not shown) and allows detection of altered receptors expressed on the cell surface using the FLAG-specific MoAb, M2. The mutant hβcΔN completely removes domains 1 and 2 (CRM1, residues E25-E232 inclusive) whereas the mutant hβcΔQP removes domains 1-3 (residues E25-N337 inclusive; the N-terminal CRM and the first domain of CRM2). hβcΔH is truncated to a site within domain 4 (residues E25-A394 deleted) and retains 45 hβc-derived, membrane proximal residues, including those that are duplicated in the constitutive hβc mutant, hβcFIΔ,[25] and the equivalent region to that retained in v-Mpl. The mutation in hβcΔWS removes residues E25-A431, including the membrane proximal WSXWS, and leaves 7 residues between the FLAG octapeptide and the transmembrane domain. These modified hβc cDNAs were cloned into the retroviral expression vector pRUFNeo.[26]

Retroviral DNA was introduced into the retroviral packaging cell line, ψ2, and stable pools of transfected ψ2 cells were produced by antibiotic selection. To test the ability f each mutant to induce factor-independent growth, w infected the murine myeloid cell line, FDC-P1, and the IL-3–dependent pro-B cell line, BaF-B03. We demonstrated surface expression f mutant recept rs n infected FDC-P1 cells by flow cytometry f ll wing staining with th M Ab, M2 (Fig 2). All truncated receptors were detectable by this procedure, indicating that receptor truncation did n t prohibit surface expression.

Proliferation of infected FDC-P1 and BaF-B03 cells. Uninfected FDC-P1 cells and G418-resistant FDC-P1 cells infected with the hβcΔN r hβcΔWS retroviruses grew in th presence of murine (m) GM-CSF but failed to proliferate after removal of growth factor. In contrast, FDC-P1 cells infected with eith r the hβcΔH r hβcΔQP

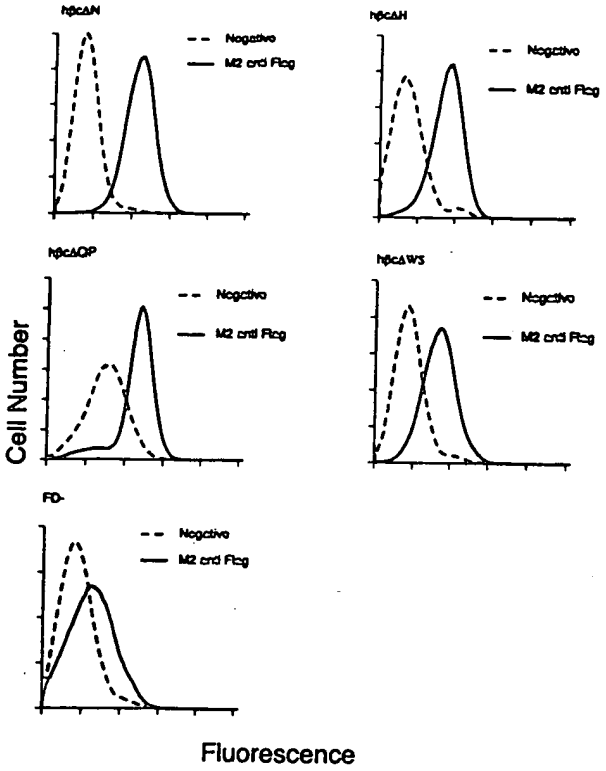


Fig 2. Expression of hβc mutants in FDC-P1 cells. Cells were stained with an irrelevant control antibody (- - -) or the anti-FLAG monoclonal antibody, M2 (——). Cell number and fluorescence are in arbitrary units with the latter plotted on a logarithmic scale.

deletion mutants grew both in the presence and absence of growth factor (Fig 3A). The different degree of proliferation observed in the absence of factor by cells infected with the hβcΔQP and hβcΔH retroviruses probably reflects the proportion of each G418-resistant population that expresses hβc. To eliminate the possibility that these mutants confer different growth rates on FDC-P1 cells, we compared proliferation over a 4-day period after selection for growth factor independence. As shown in Fig 3C, hβcΔQP- and hβcΔH-infected cells did not grow at a rate significantly different from each other, or from FDC-P1 cells grown in the presence of mGM-CSF. To show that the hβc cDNA had not undergone a rearrangement in the factor-independent FDC-P1 cells derived from infection with hβcΔQP and hβcΔH retroviruses, we recovered part of the hβc cDNA spanning the deletion, by PCR from cellular DNA with hβc specific primers. DNA from G418-resistant FDC-P1 cells infected with hβc retrovirus, and from factor-independent cells infected with pRUFhβcΔQP and pRUFhβcΔH retroviruses generated the expected products of 1,726 bp, 814 bp, and 640 bp, respectively (Fig 4), consistent with hβc retroviral insert(s) in these cells containing the appropriately mutated hβc cDNA. In common with the hβcFIΔ and hβcI374N mutants,[21] none of the truncated receptor retroviruses was capable of conferring IL-3–independent growth when used to infect the murine pro-B cell line BaF-B03 (Fig 5).

DISCUSSION

The possibility that cytokine receptors can be activated by truncation of their extracellular domain is suggested by the oncogenic v-Mpl product (see Fig 1A). An ther candidate for activation in this way is the common β chain (hβc) for IL-3, GM-CSF, and IL-5, which like Mpl, has tw repeats of the cytokine receptor motif.[33] The GM-CSF, IL-3, and IL-5 high-affinity receptors are composed of a ligand-specific α-chain (GMRα, IL3Rα, and IL5Rα), with which they f rm a low-affinity complex, and hβc.[33-36] Although hβc alone does not detectably bind cytokine it is responsible for affinity conversion and signal transduction.[34-38] Residues in the membrane proximal domain (domain 4) of hβc have been identified as critical for affinity conversion by hβc, presumably

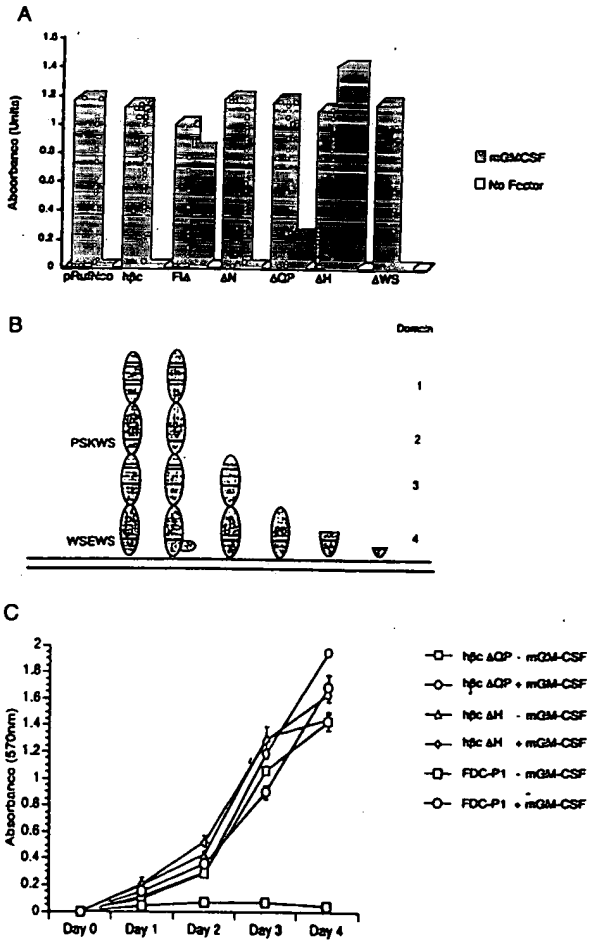


Fig 3. Proliferation of FDC-P1 cells expressing mutant forms of hβc. (A) Infected FDC-P1 cells grown in the presence (light grey) or absence (dark grey) of mGM-CSF. Proliferation assays were performed as described in Materials and Methods. FDC-P1 cells infected with pRUFNeo or with an hβcFIΔ retrovirus[25] are shown as controls. (B) A schematic representation of the hβc mutants included in the above assay. Conserved cysteine residues and WSXWS motifs are shown. hβcFIΔ contains a duplicated segment of 37 residues in domain 4.[25] (C) Growth curve comparing hβcΔQP and hβcΔH proliferation over 4 days. Cells were selected before assay by growth in factor-free medium. Proliferation of uninfected FDC-P1 cells in the presence of mGM-CSF is shown for comparison.

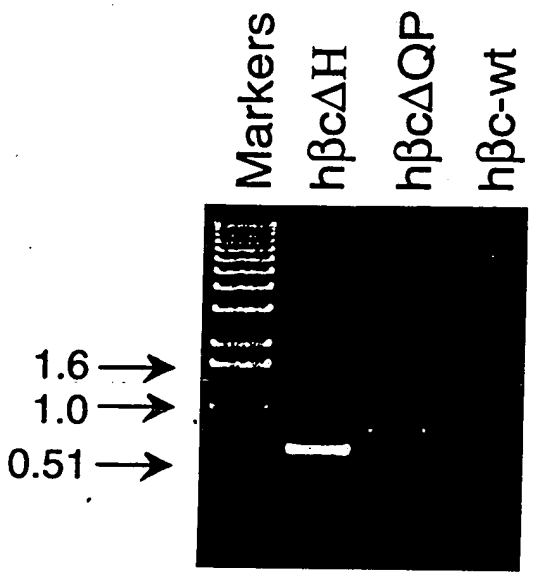


Fig 4. PCR analysis of retroviral DNA in factor-independent cells. PCR from genomic DNA to detect hβc sequences in factor-independent FDC-P1 cells derived from infection with retrovirus encoding hβcΔQP and hβcΔH (see text). PCR products were separated on a 1% agarose gel alongside molecular-weight markers (1-kb ladder; BRL, Gaithersburg, MD). The expected product sizes are 1,726 bp, 814 bp, and 640 bp for hβc, hβcΔQP, and hβcΔH, respectively. Sizes shown on the left refer to molecular weight markers.

through an interaction with residues in the N-terminal helix of the growth factor.[39] However, to date there is little information regarding the role of the first three domains of hβc in ligand binding. Several activating mutations cluster in domain 4 of hβc consistent with a critical role for this domain in activation and signaling.[21,25] Despite considerable knowledge regarding the intracellular pathways involved in signalling via hβc, there is still an incomplete understanding of the stoichiometry of the active receptor complex and the precise role of the α and β subunits in signaling.

To test the hypothesis that truncation can lead to ligand-independent cytokine receptor activation, we constructed a series of cDNAs encoding extracellular deletion mutants of hβc. Our results show that infection of the murine IL-3/GM-CSF–dependent cell line, FDC-P1, with retroviruses encoding two of these truncated forms of hβc (hβcΔQP and hβcΔH) leads to factor-independent growth. The inability of the other truncated forms of hβc (hβcΔN and hβcΔWS) to confer factor independence in FDC-P1 cells was not due to a defect in transport to the cell surface as we could detect surface expression using an MoAb directed to the introduced N-terminal FLAG epitope.

The observati n that hβcΔH confers factor independence while hβcΔWS does not, defines an extracellular region of hβc (residues H395-A431) as essential for constitutive activati n in FDC-P1 cells. It is notable that in v-Mpl, the c rresponding receptor segment is retained. The mechanism f activation of the v-*mpl* product remains unclear; h wever, based n our observations we suggest the loss of the Mpl N-terminal region is important and may expose a site of the or a heterodimer. The extracellular region retained in v-Mpl contains the WSXWS motif and the conserved β strand F′ (consensus Y/H XX R/Q VR). These motifs are also within the duplicated segment in the mutant hβcFIΔ, which previously led us to speculate that they are involved in receptor activation.[25] The conservation of this membrane proximal segment throughout the cytokine receptor family suggests that it may play an important role in the activation mechanism of class I receptors, perhaps through interactions with conserved receptor subunits. Interestingly, a previous study with chimeric receptors indicated that the extracellular domain of cytokine receptors determined, to some extent, the specific phosphorylation events that were triggered in response to ligand, suggesting interactions between extracellular sequences and signal transducers.[40]

In the growth hormone/growth hormone receptor and growth hormone/prolactin receptor complexes the corresponding segment has been shown by crystallographic studies to form part of an exposed, positively charged β-sheet in the membrane proximal domain of the ligand-bound receptor complex.[3,4] By homology the conserved residues in this segment of hβc would be positioned distal to ligand and α-subunit in a ligand-bound complex. It is possible that there are two interactive surfaces in cytokine receptor subunits that are involved in formation of oligomeric complexes. Although hβc has not been shown to form a higher order complex, we have speculated previously[21] that activation might occur through formation of $\alpha_2\beta_2$ or $\alpha\beta\gamma$ type complexes. Thus, hβc may have at least two interactive surfaces: one for association with a specific α-chain and a second that interacts with either another hβc or a third receptor component. Although there are several conserved residues between H395 and A431, it is not possible, from this work,

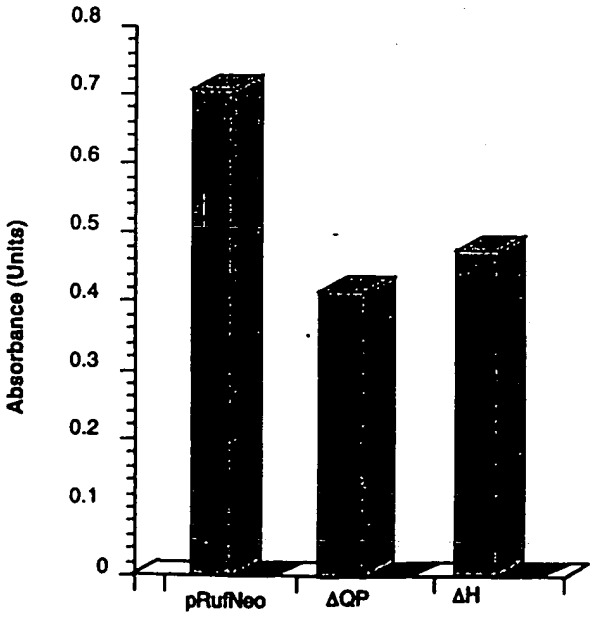


Fig 5. Proliferation of BaF-B03 cells infected with retroviruses encoding activated forms of hβc. Proliferation of infected BaF-B03 cells grown in the presence (light grey) or absence (dark grey) of murine

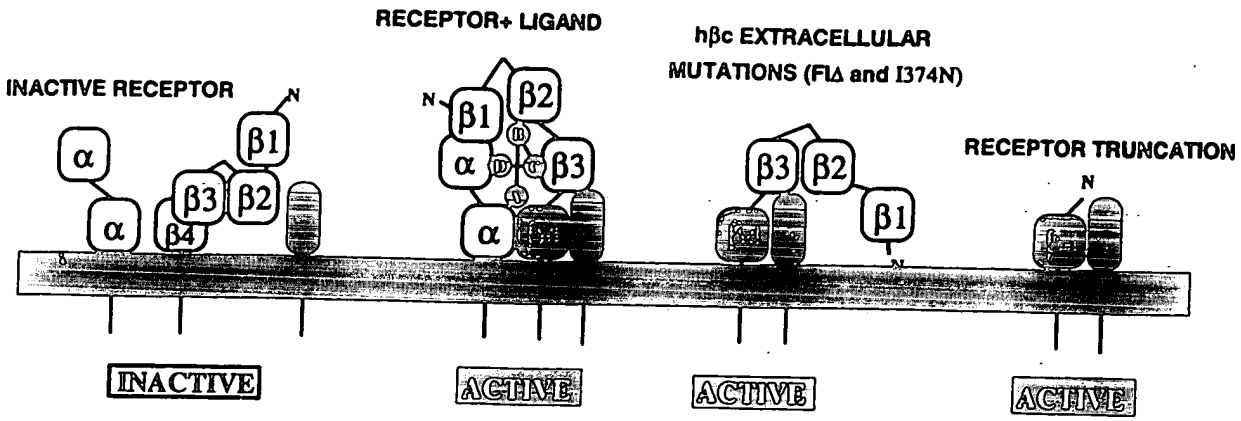


Fig 6. A model for hβc activation. One mechanism of activation of hβc may involve interaction with an unknown receptor subunit (γ) expressed in FDC-P1 cells but absent from BaF-B03 cells. Extracellular activating mutations, or extracellular truncation, are presumed to alter the receptor structure and unmask interactive residues in domain 4 leading to signaling in the absence of ligand. Ligand-mediated activation may occur through a similar process.

to define which residues are essential in receptor activation. In v-Mpl, mutation of the WSXWS motif does not impair pathogenicity of MPLV,[8] suggesting that other residues in this segment may be important.

The observation that expression of the mutant hβcΔN does not confer factor independence in FDC-P1 cells implies that the third fibronectin-like receptor domain (domain 3) may have a role in inhibiting signaling by unoccupied receptor, presumably through an interaction with residues in domain 4. Removal of this domain (as in hβcΔQP) is sufficient to lead to constitutive activation. It is tempting to speculate that the B′C′ and F′G′ loops of domain 4, which are important in ligand binding[39] (and Woodcock et al, submitted for publication), interact with ligand in such a way that the orientation of domains 3 and 4 is altered, exposing the interactive surface that allows higher order complex formation. In this way ligand may displace inhibitory interactions in the same way as truncation.

Although the mutants hβcΔQP and hβcΔH conferred factor-independence on FDC-P1 cells, they did not confer factor-independence on the IL-3–dependent pro-B cell line, BaF-B03 (Fig 5). Two other extracellular mutations, hβcFIΔ[25] and hβcI374N,[21] similarly confer factor-independent growth in FDC-P1 cells but not in Ba/F3 cells. On the other hand the hβc mutant (hβcV449E), which resembles the *neu* oncogenic receptor, and presumably leads to β-subunit dimerisation, confers IL-3–independent growth on infected BaF-B03 cells.[21] Consistent with this, chimeric receptors comprised of the GMRα or IL5Rα extracellular domain and the hβc cytoplasmic domain give a ligand-dependent response in Ba/F3 cells expressing hβc.[35,41,42] It appears from these results that hβc can be activated by homodimerization; however, the results presented here, and in the report by Jenkins et al,[21] suggest that ligand-independent f rms of hβc, with extracellular mutations, utilize an alternative mechanism of activation that may require interaction with an unknown subunit present in FDC-P1 cells but absent from BaF-B03 cells. Taken together with the results from this study, it is reasonable to suggest that the interaction of hβc with such a factor could be mediated by residues within the segment H395-A431.

In Fig 6 we present a model for hβc activation. We propose that extracellular mutations of hβc (hβcFIΔ, hβcI374N, hβcΔQP, and hβcΔH) that lead to ligand-independent activation may allow constitutive association with an unknown, myeloid-restricted, receptor component (γ) and may be mimicking a normal αβγ or βγ complex. The fact that there are several extracellular activating mutations clustered in domain 4 is consistent with activation involving the disruption of an inhibitory conformation of the unoccupied form of hβc. In this conformation we would predict that a critical structural motif within domain 4 (involving the segment H395-A431) is masked, possibly through interactions involving domain 3. We propose that the I374N substitution and the hβcFIΔ duplication can interfere with the structure of domain 4 in such a way that they lead to the disruption of inhibitory interactions.[21] By analogy with these mutants we suggest that association with ligand alters the monomeric hβc conformation and leads to association with other receptor components and the subsequent generation of intracellular signals.

ACKNOWLEDGMENT

The authors are grateful to Chris Bagley for valuable insights into receptor structures and to Alan Bishop for assistance with cell sorting. We also thank Angel Lopez, Tom Gonda, and Jo Woodcock for valuable discussions and critical reading of the manuscript.

REFERENCES

1. Bazan JF: Structural design and molecular evolution of a cytokine receptor superfamily. Proc Natl Acad Sci USA 87:6934, 1990

2. Kaczmarski RS, Mufti GJ: The cytokine receptor superfamily. Blood Rev 5:193, 1991

3. de Vos AM, Ultsch M, Kossiakoff AA: Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. Science 255:306, 1992

4. Somers W, Ultsch M, de Vos AM, Kossiakoff AA: The x-ray structure of a growth hormone-prolactin receptor complex. Nature 372:478, 1994

5. Cosman D, Lyman SD, Idzerda RL, Beckmann MP, Park LS, Goodwin RG, March CJ: A new cytokine receptor superfamily. Trends Biochem Sci 15:265, 1990

6. Fukunaga R, Ishizaka-Ikeda E, Pan CX, Seto Y, Nagata S: Functional domains of the granulocyte colony-stimulating factor receptor. EMBO J 10:2855, 1991

7. O'Neal KD, Yu Lee LY: The proline-rich motif (PRM): A novel feature of the cytokine/hematopoietin receptor superfamily. Lymphokine Cytokine Res 12:309, 1993

8. Benit L, Courtois G, Charon M, Varlet P, Dusanter Fourt I, Gisselbrecht S: Characterization of Mpl cytoplasmic domain sequences required for myeloproliferative leukemia virus pathogenicity. J Virol 68:5270, 1994

9. Ihle JN, Witthuhn BA, Quelle FW, Yamamoto K, Silvennoinen O: Signalling through the hematopoietic cytokine receptors. Annu Rev Immunol 13:369, 1995

10. Ihle J: Cytokine receptor signalling. Nature 377; 591, 1995

11. Souyri M, Vigon I, Penciolelli J-F, Heard J-M, Tambourin P, Wendling F: A putative truncated cytokine receptor gene transduced by the myeloproliferative leukaemia virus immortalises hematopoietic progenitors. Cell 63:1137, 1990

12. Kaushansky K, Lok S, Holly RD, Broudy VC, Lin N, Bailey MC, Forstrom JW, Buddle MM, Oort PJ, Hagen FS, Roth GJ, Papayannopoulou T, Foster DC: Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. Nature 369:568, 1994

13. Bartley TD, Bogenberger J, Hunt P, Li Y-S, Lu HS, Martin F, Chang M-S, Samal B, Nichol JL, Swift S, Johnson MJ, Hsu R-Y, Parker VP, Suggs S, Skrine JD, Merewether LA, Clogston C, Hsu E, Hokom MM, Hornkohl A, Choi E, Pangelinan M, Sun Y, Mar V, McNinch J, Simonet L, Jacobsen F, Xie C, Shutter J, Chute H, Basu R, Selander L, Trollinger D, Sieu L, Padilla D, Trail G, Elliot G, Izumi R, Covey T, Crouse J, Garcia A, Xu W, Del Castillo J, Biron J, Cole S, Hu MC-T, Pacifici R, Ponting I, Saris C, Wen D, Yung YP, Lin H, Bosselman RA: Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor Mpl. Cell 77:1117, 1994

14. de Sauvage FJ, Haas PE, Spencer SD, Malloy BE, Gurney AL, Spencer SA, Darbomme WC, Henzel WJ, Wong SC, Kuang WJ, Oles KS, Hultgren B, Solberg LA, Goeddel DV, Eaton DL: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl ligand. Nature 369:533, 1994

15. Lok S, Kaushansky K, Holly RD, Kuijper JL, Lofton Day CE, Oort PJ, Grant FJ, Heipel MD, Burkhead SK, Kramer JM, Bell LA, Sprecher CA, Blumberg H, Johnson R, Prunkard D, Ching AFT, Mathewes SL, Bailey MC, Forstrom JW, Buddle MM, Osborn SG, Evans SJ, Sheppard PO, Presnell SR, O'Hara PJ, Hagen FS, Roth GJ, Foster DC: Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production in vivo. Nature 369:565, 1994

16. Wendling F, Maraskovsky E, Debili N, Florindo C, Teepe M, Titeux M, Methia N, Breton Gorius J, Cosman D, Vainchenker W: cMpl ligand is a humoral regulator of megakaryocytopoiesis. Nature 369:571, 1994

17. Alexander WS, Metcalf D, Dunn AR: Point mutations within a dimer interface homology domain of c-Mpl induce constitutive receptor activity and tumorigenicity. EMBO J 14:5569, 1995

18. Watowich SS, Yoshimura A, Longmore GD, Hilton DJ, Yoshimura Y, Lodish HF: Homodimerization and constitutive activation of the erythropoietin receptor. Proc Natl Acad Sci USA 89:2140, 1992

19. Longmore GD, Lodish HF: An activating mutation in the murine erythropoietin receptor induces erythroleukaemia in mice: A cytokine receptor superfamily oncogene. Cell 67:1089, 1991

20. Watowich SS, Hilton DJ, Lodish HF: Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. Mol Cell Biol 14:3535, 1994

21. Jenkins BJ, D'Andrea R, Gonda TJ: Activating point mutations in the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors suggest the involvement of β subunit dimerisation and cell type specific molecules in signalling. EMBO J 14:4276, 1995

22. Sternberg MJE, Gullick WJ: Neu receptor dimerisation. Nature 339:587, 1989

23. Weiner DB, Liu J, Cohen JA, Williams WV, Greene MI: A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. Nature 339:230, 1989

24. Hannemann J, Hara T, Kawai M, Miyajima A, Ostertag W, Stocking C: Sequential mutations in the interleukin 3 (IL3)/GM-CSF/IL5 receptor β-subunit genes are necessary for the complete conversion to growth autonomy mediated by a truncated-β_c subunit. Mol Cell Biol 15:2402, 1995

25. D'Andrea R, Rayner J, Moretti P, Lopez A, Goodall GJ, Gonda TJ, Vadas MA: A mutation of the common receptor subunit for interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor, and IL-5 that leads to ligand independence and tumorigenicity. Blood 83:2802, 1994

26. Rayner JR, Gonda TJ: A simple and efficient procedure for generating stable expression libraries by cDNA cloning in a retroviral vector. Mol Cell Biol 14:880, 1994

27. Mann R, Mulligan RC, Baltimore D: Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33:153, 1983

28. Dexter TM, Garland J, Scott D, Scolnick E, Metcalf D: Growth of factor-dependent hemopoietic precursor cell lines. J Exp Med 152:1036, 1980

29. Hatakeyama M, Mori H, Doi T, Taniguchi T: A restricted cytoplasmic region of IL-2 receptor β chain is essential for growth signal transduction but not for ligand binding and internalisation. Cell 59:837, 1989

30. Lang RA, Metcalf D, Gough NM, Dunn AR, Gonda TJ: Expression of a haemopoietic growth factor cDNA in a factor-dependent cell line results in autonomous growth and tumorigenicity. Cell 43:531, 1985

31. Hughes SH, Payvar F, Spector D, Schimke RT, Robinson HL, Payne GS, Bishop JM, Varmus HE: Heterogeneity of genetic loci in chickens: analysis of endogenous viral and normal genes by cleavage of DNA with restriction endonucleases. Cell 18:347, 1979

32. Saiki RK: Amplification of genomic DNA, in Innes MA, Gelfand DH, Sninsky JJ, White TJ (eds): PCR Protocols: A Guide to Methods and Applications. San Diego, CA, Academic, 1990, p 13

33. Hayashida K, Kitamura T, Gorman DM, Arai K, Yokota T, Miyajima A: Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor (GM-CSF): Reconstitution of a high-affinity GM-CSF receptor. Proc Natl Acad Sci USA 87:9655, 1990

34. Kitamura T, Sato N, Arai K, Miyajima A: Expression cloning of the human IL-3 receptor cDNA reveals a shared beta subunit for the human IL-3 and GM-CSF receptors. Cell 66:1165, 1991

35. Takaki S, Murata Y, Kitamura T, Miyajima A, Tominaga A, Takatsu K: Reconstitution of the functional receptors for murine and human interleukin-5. J Exp Med 177:1523, 1993

36. Tavernier J, Devos R, Cornelis S, Tuypens T, Van der Heyden J, Fiers W, Plaetinck G: Human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific alpha chain and a beta chain shared with the receptor for GM-CSF. Cell 66:1175, 1991

37. Gearing DP, King JA, Gough NM, Nicola NA: Expression cloning of a receptor for human granulocyte-macrophage colony-stimulating factor. EMBO J 8:3667, 1989

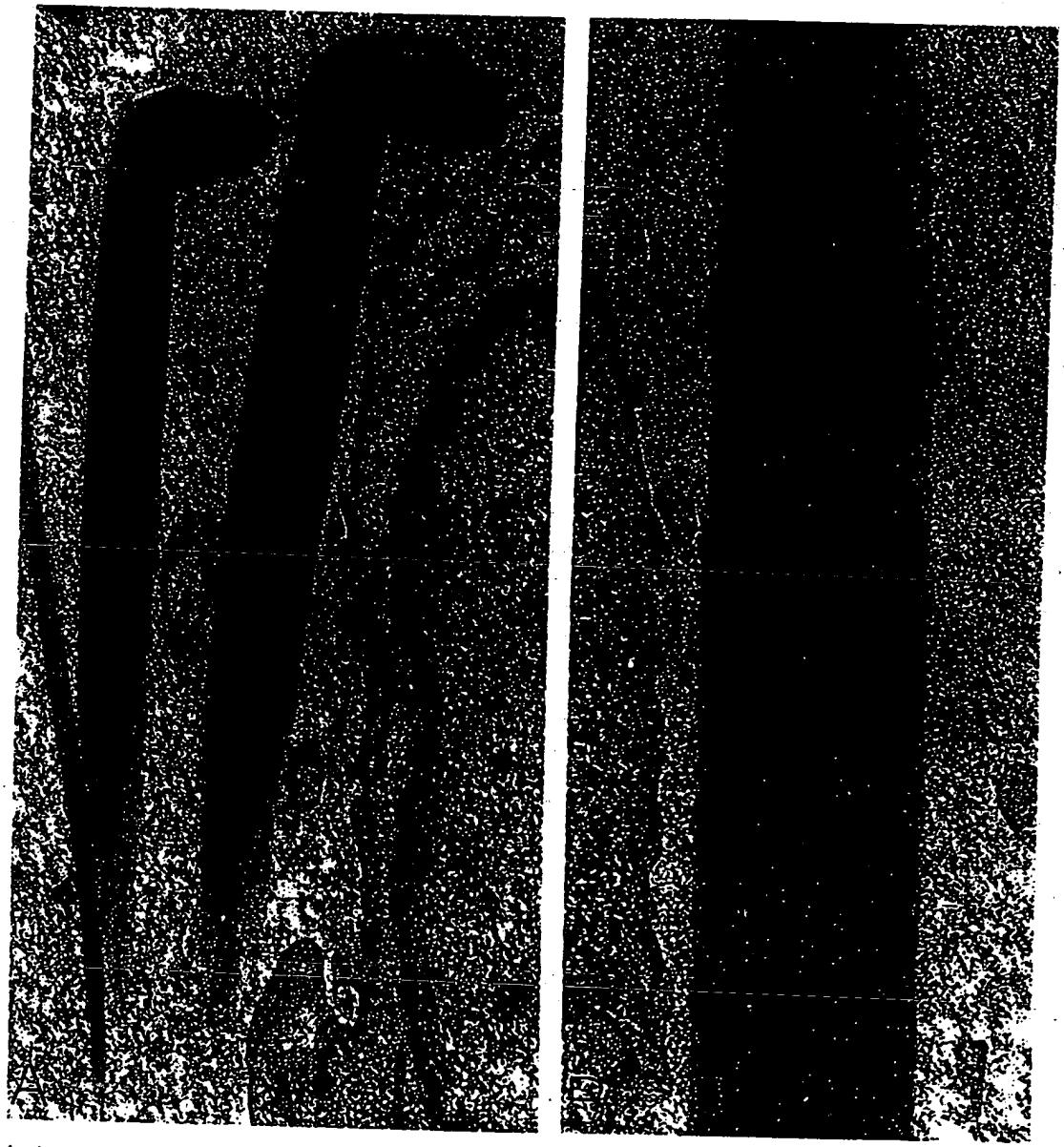
38. Murata Y, Takaki S, Migita M, Kikuchi Y, Tominaga A, Takatsu K: Molecular cloning and expression of the human interleukin 5 receptor. J Exp Med 175:341, 1992

39. Woodcock JM, Zacharakis B, Plaetinck G, Bagley CJ, Qiyu S, Hercus TR, Tavernier J, Lopez AF: Three residues in the common beta chain of the human GM-CSF, IL-3 and IL-5 receptors are essential for GM-CSF and IL-5 but not IL-3 high affinity binding and interact with Glu21 of GM-CSF. EMBO J 13:5176, 1995

40. Chiba T, Nagata Y, Machide M, Kishi A, Amanuma H, Sugiyama M, Todokoro K: Tyrosine kinase activation through the extracellular domains of cytokine receptors. Nature 362:646, 1993

41. Muto A, Watanabe S, Miyajima A, Yokota T, Arai K-I: High affinity chimeric human GM-CSF receptor carrying the cytoplasmic domain of the β subunit but not the α subunit transduces growth promoting signals in Ba/F3 cells. Biochem Biophys Res Commun 208:368, 1995

42. Eder M, Ernst TJ, Ganser A, Jubinsky PT, Inhorn R, Hoelzer D, Griffin JD: A low affinity chimeric human α/β-GM-CSF receptor induces ligand dependent proliferation in a murine cell line. J Biol Chem 269:30173, 1994



Fibrin in the peripheral blood of a patient with acute myelocytic leukemia (A). Note the periodicity of fibrin at higher magnification (B). (Courtesy of Ann M. Dvorak, MD, Department of Pathology, Beth Israel Hospital, Harvard Medical School, Boston, MA 02215.)

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 11168–11172, November 1994
Biochemistry

Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions

(zinc finger–DNA interaction/recognition code/binding-site signature/protein design/DNA–protein interaction)

YEN CHOO AND AARON KLUG

Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, United Kingdom

Contributed by Aaron Klug, June 6, 1994

ABSTRACT In the preceding paper [Choo, Y. & Klug, A. (1994) Proc. Natl. Acad. Sci. USA 91, 11163–11167], we showed how selections from a library of zinc fingers displayed on phage yielded fingers able to bind to a number of DNA triplets. Here, we describe a technique to deal efficiently with the converse problem—namely, the selection of a DNA binding site for a given zinc finger. This is done by screening against libraries of DNA triplet binding sites randomized in two positions but having one base fixed in the third position. The technique is applied here to determine the specificity of fingers previously selected by phage display. We find that some of these fingers are able to specify a unique base in each position of the cognate triplet. This is further illustrated by examples of fingers which can discriminate between closely related triplets as measured by their respective equilibrium dissociation constants. Comparing the amino acid sequences of fingers which specify a particular base in a triplet, we infer that in most instances, sequence-specific binding of zinc fingers to DNA can be achieved by using a small set of amino acid–nucleotide base contacts amenable to a code.

In principle, rules governing protein–DNA interactions can be deduced from a large database of correlations between the amino acid sequences of the proteins and the nucleotide sequences of their optimal binding sites. To this end, we have shown in the preceding paper (1) that functionally equivalent zinc fingers which bind to a given DNA sequence can be selected from a phage display library. However, determination of the optimal binding site for these fingers is still required, as a safeguard against spurious selections. One can determine the optimal binding sites of these (and other) proteins, by selection from libraries of randomized DNA. This approach, the principle of which is essentially the converse of zinc finger phage display, would provide an equally informative database from which the same rules can be independently deduced. However, until now the favored method for binding-site determination, involving iterative selection and amplification of target DNA followed by sequencing, has been a laborious process not conveniently applicable to the analysis of a large database (2, 3).

We present here a convenient and rapid method which can reveal the optimal binding site(s) of a DNA-binding protein by single-step selection from small libraries, and use this to check the binding-site preferences of those zinc fingers selected previously by phage display (1). For this application, we use 12 different minilibraries of the binding site for transcription factor Zif268, each one with the central triplet having one position defined with a particular base pair and the other two positions randomized. Each library therefore comprises 16 oligonucleotides and offers a number of potential binding sites to the middle finger, provided that the latter can tolerate the defined base pair. Each zinc finger phage is screened against all 12 libraries individually immobilized in wells of a microtiter plate, and binding is detected by an enzyme immunoassay. Thus, a pattern of acceptable bases at each position is disclosed, which we call a binding-site signature. The information contained in a binding-site signature encompasses the repertoire of binding sites recognized by a zinc finger.

The binding-site signatures obtained by using zinc finger phage selected as described in the preceding paper (1) reveal that the selection has yielded some highly sequence-specific zinc fingers which discriminate at all three positions of a triplet. From measurements of equilibrium dissociation constants, we find that these fingers bind tightly to the triplets indicated in their signatures and discriminate against closely related sites usually by at least a factor of 10. The binding-site signatures allow us to infer rules for a specificity code for the interactions of zinc fingers with DNA.

MATERIALS AND METHODS

Binding-Site Signatures. Flexible flat-bottomed 96-well plates (Falcon) were coated overnight at 4°C with streptavidin (0.1 mg/ml in 0.1 M $NaHCO_3$, pH 8.6/0.03% NaN_3). Wells were blocked by incubation for 1 hr with PBS/Zn (phosphate-buffered saline plus 50 μM zinc acetate) containing 2% (wt/vol) fat-free dried milk (Marvel) and were washed three times with PBS/Zn containing 0.1% Tween and three times with PBS/Zn. The "bound" strand of each oligonucleotide library was made synthetically and the other strand was extended from a 5'-biotinylated universal primer by DNA polymerase I (Klenow fragment). Products of fill-in reactions were added to wells (0.8 pmol of DNA library in each) in PBS/Zn for 15 min and then washed once with PBS/Zn containing 0.1% Tween and once with PBS/Zn. Overnight bacterial cultures each containing a selected zinc finger phage (1) were grown at 30°C in 2×TY medium containing 50 μM zinc acetate and 15 μg of tetracycline per ml (2×TY/Zn/Tet). Culture supernatants containing phage were diluted 10-fold by addition of PBS/Zn containing 2% (wt/vol) fat-free dried milk, 1% (vol/vol) Tween 20 and 20 μg of sonicated salmon sperm DNA per ml. Diluted phage solutions (50 μl) were applied to wells and binding was allowed to proceed for 1 hr at 20°C. Unbound phage were removed by washing five times with PBS/Zn containing 1% Tween and then three times with PBS/Zn. Bound phage were detected as described (4) or by using horseradish peroxidase-conjugated anti-M13 IgG (Pharmacia) and quantitated with SOFTMAX 2.32 (Molecular Devices).

Determination of Apparent Equilibrium Dissociation Constants (K_d Values). Overnight bacterial cultures were grown in 2×TY/Zn/Tet at 30°C. Culture supernatants containing phage were diluted 2-fold by the addition of PBS/Zn containing 4% fat-free dried milk, 2% Tween 20, and 40 μg of sonicated salmon sperm DNA per ml. Binding reaction mixtures containing appropriate concentrations of specific

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

5′-biotinylated DNA and equal volumes of zinc-finger phage solution were allowed to equilibrate for 1 hr at 20°C. All DNA was captured on streptavidin-coated paramagnetic beads (500 μg per well), which were subsequently washed six times with PBS/Zn containing 1% Tween and then three times with PBS/Zn. Bound phage were detected with horseradish peroxidase-conjugated anti-M13 IgG (Pharmacia) and developed as described (4). Optical densities were quantitated with SOFTMAX 2.32 (Molecular Devices).

K_d values were estimated by fitting to the equation K_d = [DNA][protein]/[DNA–protein] with the program KALEIDAGRAPH version 2.0 (Synergy Software, Reading, PA). Owing to the sensitivity of the ELISA used to detect protein–DNA complexes, we can use zinc finger phage concentrations far below those of the DNA, as is required for accurate calculations of K_d. The technique we use has the advantage that while the concentration of DNA (variable) must be known accurately, that of the zinc fingers (constant) need not be known (5). This circumvents the problem of calculating the number of zinc finger peptides expressed on the tip of each phage, although since only 10–20% of the gene III protein (pIII) carries such peptides, we would expect on average less than one copy per phage. Binding is performed in solution to prevent any effects caused by the avidity (6) of phage for DNA immobilized on a surface. Moreover, in this case measurements of K_d by ELISA are made possible because

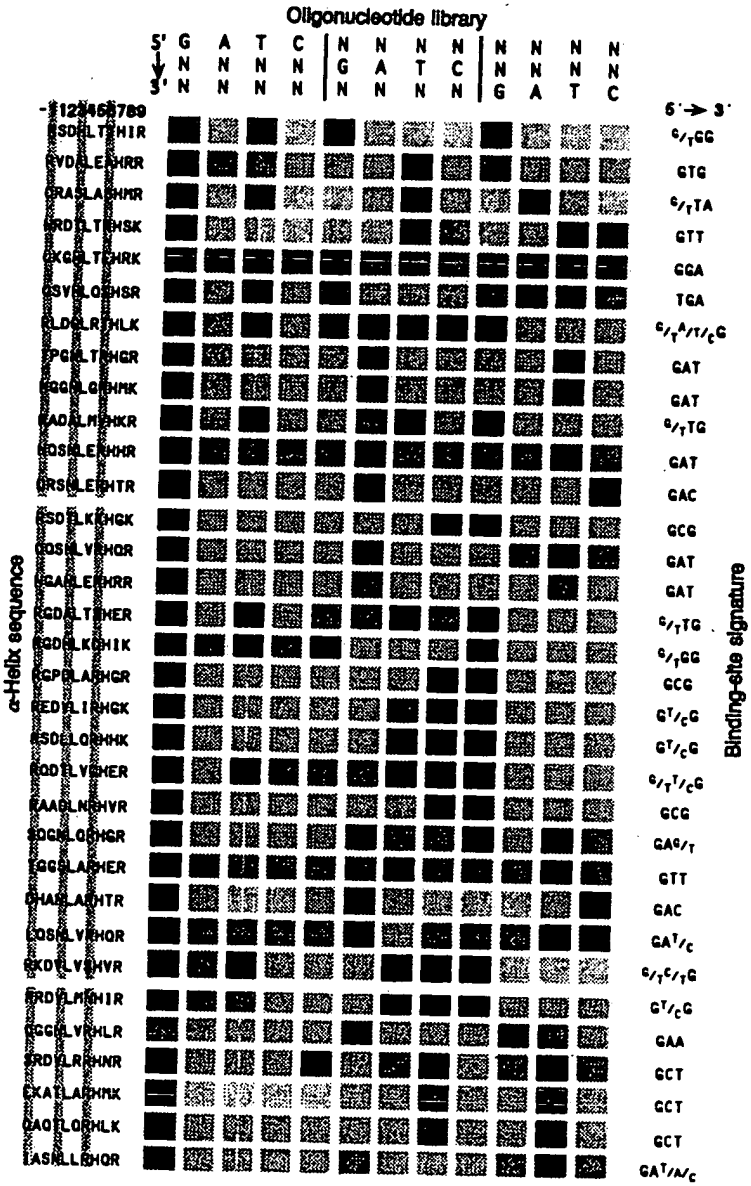


FIG. 1. Binding-site signatures of individual zinc finger phage. The diagram is of raw data and represents binding of zinc finger phage to randomized DNA immobilized in the wells of microtiter plates. To test each zinc finger phage against each oligonucleotide library (see text), DNA libraries are applied to columns of wells (down the plate), while rows of wells (across the plate) contain equal volumes of a solution of a zinc finger phage. The identity of each library is given as the middle triplet of the "bound" strand of Zif268 operator, where N represents a mixture of all four nucleotides. The zinc finger phage is specified by the sequence of the variable region of the middle finger, numbered relative to the first helical residue (position +1), and the three primary recognition positions are highlighted. Bound phage are detected by an enzyme immunoassay. The approximate strength of binding is indicated by a gray scale proportional to the enzyme activity. From the pattern of binding to DNA libraries, called the signature of each clone, one or a small number of binding sites can be read off; these are shown at right.

equilibrium is reached in solution prior to capture n the solid phase.

RESULTS AND DISCUSSION

Binding-Site Signature of the Second Zinc Finger of Zif268. The top row of Fig. 1 shows the signature of the second finger of wild-type Zif268. From the pattern of strong signals indicating binding to oligonucleotide libraries having GNN, TNN, NGN, and NNG as the middle triplet, it emerges that the optimal binding site for this finger is (T/G)GG, in accord with the published consensus sequence (7). This has implications for the interpretation of the x-ray crystal structure of Zif268 solved in complex with a consensus operator having TGG as the middle triplet (8). For instance, His at position +3 of the middle finger was modeled as donating a hydrogen bond to N7 of guanine, suggesting an equivalent contact to be possible with N7 of adenine, but from the binding-site signature we can see that there is discrimination against adenine. This implies that the His may prefer to make a hydrogen bond to O6 of guanine or a bifurcated hydrogen bond to both O6 and N7 or that a steric clash with the amino group of adenine may prevent a tight interaction with this base. Thus, from consideration of the stereochemistry of double-helical DNA, binding-site signatures can give insight into the details of zinc finger–DNA interactions.

Amino Acid–Nucleotide Base Contacts in Zinc Finger–DNA Complexes Deduced from Binding-Site Signatures. The binding-site signatures of other zinc fingers (Fig. 1) reveal that the phage selections we performed in our previous study (1) have yielded highly sequence-specific DNA-binding proteins. Some of these are able to specify a unique sequence for the middle triplet of a variant Zif268 binding site and are therefore more specific than is Zif268 itself for its consensus site. Moreover, one can identify the fingers which recognize a particular oligonucleotide library—that is to say a specific base at a defined position—by looking down the columns of Fig. 1. By comparing the amino acid sequences of these fingers we can identify any residues which have genuine preferences for particular bases on bound DNA. With a few exceptions, these are as previously predicted on the basis of phage display (1) and are summarized in Fig. 2.

The binding-site signatures also reveal an important feature of our phage display library which is crucial to the interpretation of our selection results. All the fingers in our panel, regardless of the amino acid present at position +6, are able to recognize guanine or both guanine and thymine at the 5′ end of a triplet. Our explanation for this is that the 5′ position of the middle triplet is fixed as either guanine or thymine by a contact from the invariant Asp at position +2 of finger 3 to the partner of either base on the complementary strand, analogous to those seen in the Zif268 (8) and tramtrack (9) crystal structures (a contact to the NH_2 of cytosine or adenine, respectively, in the major groove). Therefore Asp at position +2 of finger 3 is dominant over the amino acid present at position +6 of the middle finger, precluding the possibility of recognition of adenine or cytosine at the 5′ position. Future libraries must be designed with this interaction omitted or the position varied. Interestingly, given the framework of the conserved regions of the three fingers, we can identify a rule in the second finger which specifies a frequent interaction with both guanine and thymine—namely, the occurrence of Ser or Thr at position +6, which may donate a hydrogen bond to either base.

Modulation of Base Recognition by Auxiliary Positions. As we have n ted above, positi n +2 is able to specify the base directly 3′ of the "cognate triplet" and can thus work in conjunction with position +6 of the preceding finger. The binding-site signatures, while pointing to amino acid–base contacts from the three primary positions, indicate that

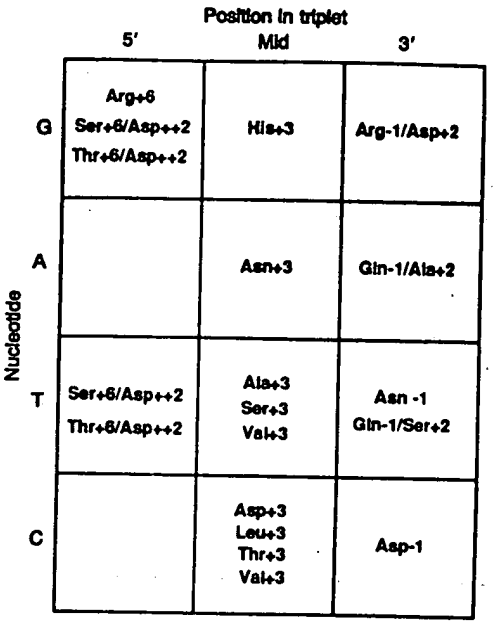


FIG. 2. Summary of frequently observed amino acid–nucleotide base contacts in interactions of selected zinc fingers with DNA. The given contacts comprise a "syllabic" recognition code (see text) for appropriate triplets. Cognate amino acids and their positions in the α-helix are entered in a matrix relating each base to each position of a triplet. Auxiliary amino acids from position +2 can enhance or modulate specificity of amino acids at position −1, and these are listed as pairs. Ser or Thr at position +6 permit Asp at +2 of the following finger (denoted Asp++2) to specify both guanine (G) and thymine (T) indirectly, and the pairs are listed. The specificity of Ser at +3 for T and Thr at +3 for cytosine (C) may be interchangeable in rare instances, whereas Val at +3 appears to be consistently ambiguous.

auxiliary positions can play other parts in base recognition. A clear case in point is Gln at position −1, which is specific for adenine at the 3′ end of a triplet when position +2 is a small nonpolar amino acid such as Ala but is specific for thymine when a polar residue such as Ser is at position +2. The strong correlation between Arg at position −1 and Asp at position +2, the basis of which is understood from the x-ray crystal structures of zinc fingers (8, 9), is another instance of interplay between these two positions. Thus the amino acid at position +2 is able to modulate or enhance the specificity of the amino acid at other positions.

At position +3, a different type of modulation is seen in the case of Thr and Val, which most often prefer cytosine in the middle position of a triplet, but in some zinc fingers are able to recognize both cytosine and thymine. This ambiguity occurs possibly as a result of different hydrophobic interactions involving the methyl groups of these residues, and here a flexibility in the inclination of the finger rather than an effect from another position *per se* may be the cause of ambiguous reading.

Quantitative Measurements of Dissociation Constants. The binding-site signature of a zinc finger reveals its differential base preferences at a given concentration of DNA. As the concentration of DNA is altered, one can expect the binding site signature of any clone to change, being more distinctive at l w [DNA], and becoming less so at higher [DNA] as the K_d of less fav rable sites is approached and further bases become acceptable at each position f the triplet. Further, because two base positions are randomly occupied in any ne library f oligonucleotides, binding-site signatures are not formally able to exclude the possibility f context depen-

dence for some interactions. Therefore to supplement binding-site signatures, which are essentially comparative, quantitative determinations of the K_d values of each phage for different DNA binding sites are required. After phage display selection and binding-site signatures, this is the third and definitive stage in assessing the specificity of zinc fingers.

Examples of such studies presented in Fig. 3 reveal that zinc finger phages bind the operators indicated in their binding-site signatures with K_d values in the range of 10^{-8} to 10^{-9} M and can discriminate against closely related binding sites by factors greater than an order of magnitude. Indeed, Fig. 3 shows such differences in affinity for binding sites which differ in only one out of nine base pairs. Since the zinc fingers in our panel were selected from a library by noncompetitive affinity purification, there is the possibility that fingers which are even more discriminatory can be isolated by a competitive selection process.

Measurements of K_d allow different triplets to be ranked in order of preference according to the strength of binding. The examples here indicate that the contacts from either position −1 or +3 can contribute to discrimination. Also, the ambiguity in certain binding-site signatures referred to above can be shown to have a basis in the equal affinity of certain fingers for closely related triplets. This is demonstrated by the K_d values of the finger containing the amino acid sequence RGDALTSHER for the triplets TTG and GTG.

A Code for Zinc Finger–DNA Recognition. One would expect that the versatility of the zinc finger motif will have allowed evolution to develop various modes of binding to DNA (and even to RNA) which will be too diverse to fall under the scope of a single code. However, although a code may not apply to all zinc finger–DNA interactions, there is now convincing evidence that a code applies to a substantial subset. This code will fall short of being able to predict unfailingly the DNA binding-site preference of any given zinc finger from its amino acid sequence but may yet be sufficiently comprehensive to allow the design of zinc fingers with specificity for a given DNA sequence.

Using the selection methods of phage display (1) and of binding-site signatures, we find that in the case of Zif268-like zinc fingers, DNA recognition involves four fixed principal (three primary and one auxiliary) positions on the α-helix, from which a limited and specific set of amino acid–base contacts result in recognition of a variety of DNA triplets. In other words, a code can describe the interactions of zinc fingers with DNA. Toward this code, we can propose amino acid–base contacts for almost all the entries in a matrix relating each base to each position of a triplet (Fig. 2). Where there is overlap, our results complement those of Desjarlais and Berg (10, 11), who have derived similar rules by altering zinc finger specificity, using database-guided mutagenesis.

Combinatorial Use of the Coded Contacts. The individual base contacts listed in Fig. 2, though part of a code, may not always result in sequence-specific binding to the expected base triplet when used in any combination. First, we must be aware of the possibility that zinc fingers may not be able to recognize certain combinations of bases in some triplets by use of this code, or even at all. Otherwise, the majority of inconsistencies may be accounted for by considering variations in the inclination of the trident reading head of a zinc finger with respect to the triplet with which it is interacting. It appears that the identity of an amino acid at any one α-helical position is attuned to the identity of the residues at the other two positions to allow three base contacts to occur simultaneously. Therefore, for example, in order that Ala may pick out thymine in the triplet GTG, Arg must not be used to recognize guanine from position +6, since this would distance the Ala residue too far from the DNA (see for example the finger containing the amino acid sequence RGDALTSHER). Second, since the pitch of the α-helix is 3.6

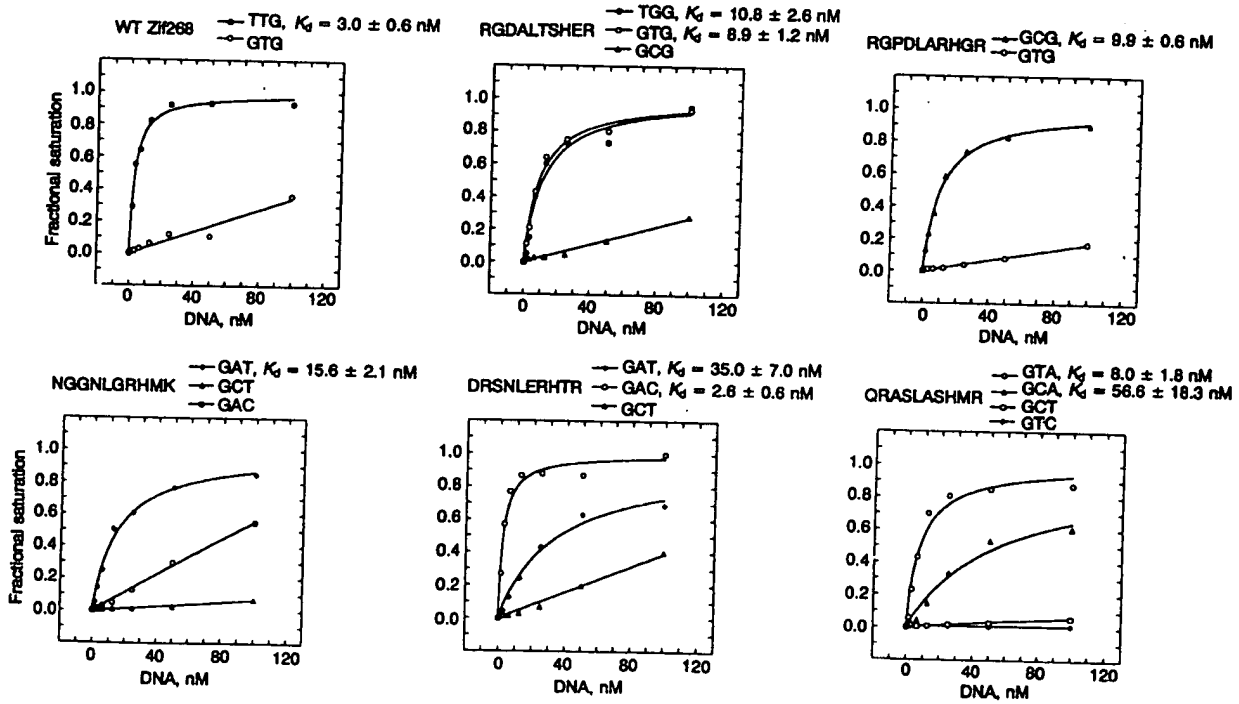


FIG. 3. Determination of apparent equilibrium dissociation constants of zinc finger phage for variants of the Zif268 binding site, showing discrimination of closely related triplets by the middle finger, usually by factors of >10. The two outer fingers carry the native sequence, as do the two cognate outer DNA triplets. The sequence of amino acids occupying helical positions −1 to +9 of the varied middle finger is shown in each case. WT, wild type (RSDHLTTHIR).

amino acids per turn, positions −1, +3, and +6 are not an integral number f turns apart, so that positi n +3 is nearer to the DNA than is −1 r +6. Hence, for example, sh rt amin acids such as His and Asn, rather than the longer Arg and Gln, are used f r the recognition of purines in the middle position of a triplet.

As a consequence of these distance effects, we might say that the code is not really "alphabetic" (always identical amino acid–base contact) but rather "syllabic" (use of a small repertoire of amino acid–base contacts). An alphabetic code would involve only four rules, but syllabicity adds an additional level of complexity, since systematic combinations of rules comprise the code. Nevertheless, the recognition of each triplet is still best described by a code of syllables, rather than a catalogue of "logograms" (idiosyncratic amino acid–base contact depending on triplet).

Conclusions. The syllabic code of interactions with DNA is made possible by the versatile framework of the zinc finger: this allows an adaptability at the interface with DNA by slight changes of orientation, which in turn maintains a stoichiometry of one coplanar amino acid per base pair in many different complexes. Given this mode of interaction between amino acids and bases, it is to be expected that recognition of guanine and adenine by Arg and Asn/Gln, respectively, is an important feature of the code; but remarkably, other interactions can be more discriminatory than was anticipated (12). Conversely, it is clear that degeneracy can be programmed in the zinc fingers in varying degrees, allowing for intricate interactions with different regulatory DNA sequences (7, 13). One can see how this principle makes possible the regulation of differential gene expression by a limited set of transcription factors.

As we have noted, the versatility of the finger motif will most likely allow other modes of binding to DNA. Similarly, we must take into account the malleability of nucleic acids, such as was observed in ref. 9, where a deformation of the double helix at a flexible base step allows a direct contact from Ser at position +2 of finger 1 to a thymine at the 3′ position of the cognate triplet. Even in our selections there are instances of fingers whose binding mode is obscure and may require structural analyses for clarification. Thus, water may be seen to play an important role, for example, where short side chains such as those of Asp, Asn, or Ser interact with bases from position −1 (14, 15).

Eventually, it might be possible to develop a number of codes describing zinc finger binding to DNA, which could predict the binding-site preferences of some zinc fingers from their amino acid sequences. The functional amino acids selected in this study at positions −1, +3, and, to some extent, +6 are very frequently observed at the same positions in naturally occurring fingers (e.g., see figure 4 of ref. 16), supporting the existence of coded contacts from these three positions. However, the lack of definitive predictive methods is not a serious practical limitation, as current laboratory techniques (this paper and refs. 2 and 3) will allow the identification of binding sites for a given DNA-binding protein. Rather, we can apply phage selection and a knowledge of the recognition rules to the converse problem, the design of proteins to bind predetermined DNA sites.

Prospects for the Design of DNA-Binding Proteins. The ability to manipulate the sequence specificity f zinc fingers implies that we are on the eve of designing DNA-binding proteins with desired specificity for applications in medicine and research (11, 17). This is possible because of the modular nature of th zinc finger, by contrast to all other DNA-binding motifs, since DNA sites can be recognized by appropriate combinations of independently acting fingers linked in tandem.

The coded interactions of zinc fingers with DNA can be used to model the specificity of individual zinc fingers *de novo* or, more likely, in conjunction with phage display selection of suitable candidates. In this way, according to requirements, one could modulate the affinity for a given binding site or even engineer an appropriate degree of indiscrimination at particular base positions. Moreover, the additive effect of multiply repeated domains offers the opportunity to bind specifically and tightly to extended, and hence very rare, genomic loci. Thus, zinc finger proteins might well be a good alternative to the use of antisense nucleic acids in suppressing or modifying the action of a given gene, whether normal or mutant. To this end, extra functions could be introduced into these DNA-binding domains by appending suitable natural or synthetic effectors.

We thank L. Fairall, A. Griffiths, D. Rhodes, and J. Schwabe for critical reading of the manuscript. Y.C. thanks the Medical Research Council and the British Council for funding.

1. Choo, Y. & Klug, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11163–11167.
2. Thiesen, H. J. & Bach, C. (1990) *Nucleic Acids Res.* **18**, 3203–3209.
3. Pollock, R. & Treisman, R. (1990) *Nucleic Acids Res.* **18**, 6197–6204.
4. Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I., Waterhouse, P., Crosby, W. L., Kontermann, R. E., Jones, P. T., Low, N. M., Allison, T. J., Prospero, T. D., Hoogenboom, H. R., Nissim, A., Cox, J. P. L., Harrison, J. L., Zaccolo, M., Gherardi, E. & Winter, G. (1994) *EMBO J.* **13**, 3245–3260.
5. Choo, Y. & Klug, A. (1993) *Nucleic Acids Res.* **21**, 3341–3346.
6. Marks, J. D., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1992) *J. Biol. Chem.* **267**, 16007–16010.
7. Christy, B. & Nathans, D. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8737–8741.
8. Pavletich, N. P. & Pabo, C. O. (1991) *Science* **252**, 809–817.
9. Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T. & Rhodes, D. (1993) *Nature (London)* **366**, 483–487.
10. Desjarlais, J. R. & Berg, J. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 7345–7349.
11. Desjarlais, J. R. & Berg, J. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2256–2260.
12. Seeman, N. D., Rosenberg, J. M. & Rich, A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 804–808.
13. Harrison, S. D. & Travers, A. A. (1990) *EMBO J.* **9**, 207–216.
14. Qian, Y. Q., Otting, G. & Wuthrich, K. (1993) *J. Am. Chem. Soc.* **115**, 1189–1190.
15. Shakked, Z., Guzikevich-Guerstein, G., Frolow, F., Rabinovich, D., Joachimiak, A. & Sigler, P. B. (1994) *Nature (London)* **368**, 469–473.
16. Desjarlais, J. R. & Berg, J. M. (1992) *Proteins* **12**, 101–104.
17. Rebar, E. J. & Pabo, C. O. (1994) *Science* **263**, 671–673.

is distinct rigid body movement with respect to the B pentamer. Residues A:1–222 (that is, the A1 fragment plus the long N-terminal helix of A2) move by a 5° rotation, with residues 216–222 functioning as a hinge (Fig. 3*a*). This type of rotation is also found in other crystal f rms f LT without bound sugars (ur unpublished results), but the remarkable flexibility may be important for translocati n f LT and cholera toxin across the membrane.

One unexpected result was the well-defined density for five C-terminal residues of the A2 fragment (comprising residues A:196–240), which were poorly defined in the native map at 2.3-Å resolution[25]. In the LT–lactose complex, residues A2:232–235 adopt a helical conformation (Fig. 3*b*). Hence the unusual A2 fragment consists of (1) a long N-terminal helix, interacting with the A1 enzyme; and (2) an extended chain, comprising residues 227–231, protruding through the central pore of the B pentamer along the 5-fold axis; and (3) a one-and-a-half turn C-terminal helix (Fig. 3*c*). Temperature factors are relatively high for residues 232–235 (on average 45 Å^2 versus an average of 27 Å^2 for the entire A2, and 18 Å^2 for the B pentamer) and hence this helix may be poised for conformational change. The C-terminal helix is also found in a higher-resolution native LT structure based on 1.95 Å data (T.K.S. *et al.*, manuscript in preparation) and therefore seems to be an intrinsic feature of LT.

The orientation of the bound galactose (Fig. 3*a*) is such that the C1 linker atom is in the middle of the 'convoluted' surface of the B pentamer, about 8 Å horizontally away from the side of the pentamer and at least 25 Å vertically away from the opposite 'flat' or 'A-binding' surface. This distance is too far for the remaining sugar residues of the G_{M1} oligosaccharide to span because they extend 25 Å at most from the hydrophobic part of the membrane[26]. Binding with the A subunit facing the membrane would not only force the A subunit into the membrane but also insert the hydrophilic B pentamer into the hydrophobic part of the membrane by at least 8 Å, in contradiction to biochemical data[10,11,18]. As there is little or no evidence for any large conformational change in the B pentamer upon G_{M1} binding[9,14,20], the initial binding of LT or cholera toxin to G_{M1}-containing membranes most probably occurs with the A subunit pointing away from the membrane, when all five ganglioside binding sites are occupied. This mode of binding has already been proposed[5,10] but others have come to different conclusions[11,12,18]. It should be noted that here the C terminus of A2 interacts with the membrane (Fig. 3), in good agreement with photoaffinity labelling studies[10]. The C-terminal sequence of A2 (ArgAspGluLeu) closely resembles the LysAspGluLeu sequence in cholera toxin; this KDEL sequence may act as a retention signal in the endoplasmic reticulum membrane[27] and could be important for the interaction of both toxins with the cell membrane.

Finally, it has been pointed out to us that the topology of monomeric *Staphylococcus aureus* nuclease[28] is essentially the same as that of the B subunit of LT, although the relative orientation of the secondary-structure elements is different (A. Muzzin, personal communication). But the position of the galactose-binding site described here corresponds roughly with the active site of the nuclease, which may be of interest from the point of view of evolution in that LT, cholera toxin and the nuclease are all secreted by bacterial pathogens. □

Received 9 September; accepted 30 October 1991.

1. Eidels, L., Proia, R. L. & Hart, D. A. *Microbiol. Rev.* **47**, 596–620 (1983).
2. van Heyningen, S. in *Current Topics in Membranes and Transport* Vol. 18, 445–471 (Academic, New York, 1983).
3. Moss, J. & Vaughan, M. *Adv. Enzym.* **61**, 303–379 (1988).
4. Montesano, R., Roth, J., Robert, A. & Orci, L. *Nature* **296**, 651–653 (1982).
5. Janicot, M., Fouque, F. & Desbuquois, B. *J. biol. Chem.* **266**, 12858–12865 (1991).
6. Finkelstein, R. A. in *Immunochemical and Molecular Genetic Analysis of Bacterial Pathogens* (eds Owen, P. & Foster, T. J.) 85–102 (Elsevier, New York, 1988).
7. Sattler, J., Schwarzmann, G., Staerk, J., Ziegler, W. & Wiegandt, H. *Hoppe Seyler's Z. physiol. Chem.* **358**, 159–163 (1977).
8. Clements, J. D. & Finkelstein, R. A. *Infect. Immunity* **24**, 760–769 (1979).
9. De Wolf, M. J. S., Fridkin, M. & Kohn, L. D. *J. biol. Chem.* **256**, 5489–5496 (1981).
10. Tomasi, M. & Montecucco, C. *J. biol. Chem.* **256**, 11177–11181 (1981).
11. Wisnieski, B. J. & Bramhall, J. S. *Nature* **289**, 319–321 (1981).
12. Dwyer, J. D. & Bloomfield, V. A. *Biochemistry* **21**, 3227–3231 (1982).
13. Holmgren, J., Fredman, P. Lindblad, M., Svennerholm, A.-M. & Svennerholm, L. *Infect. Immu* **36**, 424–433 (1982).
14. Ludwig, D. S., Holmes, R. K. & Schoolnik, G. K. *J. biol. Chem.* **260**, 12528–12534 (1985).
15. Tsuji, T., Honda, T., Miwatani, T., Wakabayashi, S. & Matsubara, H. *J. biol. Chem.* **260**, 8552–8 (1985).
16. Griffiths, S. L., Finkelstein, R. A. & Critchley, D. R. *Biochem. J.* **238**, 313–322 (1986).
17. Fukuta, S., Magnani, J. L., Twiddy, E. M., Holmes, R. K. & Ginsburg, V. *Infect. Immunity* **56**, 1748–1 (1988).
18. Ribi, H. A., Ludwig, D. S., Mercer, K. L., Schoolnik, G. K. & Kornberg, R. D. *Science* **239**, 1272–1 (1988).
19. Goins, B. & Freire, E. *Biochemistry* **27**, 2046–2052 (1988).
20. Surewicz, W. K., Leddy, J. J. & Mantsch, H. H. *Biochemistry* **29**, 8106–8111 (1990).
21. Svennerholm, A.-M., Holmgren, J., Sack, D. A. & Bardhan, P. K. *Lancet* **i**, 305–307 (1982).
22. de Aizpurua, H. J. & Russell-Jones, G. J. *J. exp. Med.* **167**, 440–451 (1988).
23. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. *J. biol. Chem.* **266**, 5226–5237 (1991).
24. Rutenber, E. & Robertus, J. D. *Proteins* **10**, 260–269 (1991).
25. Sixma, T. K. *et al. Nature* **351**, 371–377 (1991).
26. McDaniel, R. V. & McIntosh, T. J. *Biophys. J.* **49**, 94–96 (1986).
27. Pelham, H. R. B. *EMBO J.* **8**, 3171–3176 (1989).
28. Arnone, A. *et al. J. biol. Chem.* **246**, 2302–2316 (1971).
29. Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
30. Leong, J., Vinal, A. C. & Dallas, W. S. *Infect. Immunity* **48**, 73–77 (1985).
31. Pronk, S. E. *et al. J. biol. Chem.* **260**, 13580–13584 (1985).
32. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. *Acta crystallogr.* **A43**, 489–501 (1987).
33. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
34. Lesk, A. M. & Hardman, K. D. *Meth. Enzym.* **115**, 381–390 (1985).
35. Priestle, J. P. *J. appl. Crystallogr.* **21**, 572–576 (1988).
36. Sanchez, J., Hirst, T. R. & Uhlin, B. E. *Gene* **64**, 265–275 (1988).

ACKNOWLEDGEMENTS. We thank B. Witholt and J. Kingma for purified LT, and C. Verlinde, A. Mu and T. Hirst for discussion. This project has been sponsored by the Netherlands Foundation Chemical Research (S.O.N.) and the Netherlands Organisation for Scientific Research (N.W.O.)

Selection of single-strand d DNA molecules that bind and inhibit human thrombin

Louis C. Bock, Linda C. Griffin, John A. Latham, Eric H. Vermaas & John J. Toole*

Gilead Sciences Inc., 346 Lakeside Drive, Foster City, California 94404, USA

APTAMERS[1] are double-stranded DNA or single-stranded RN molecules that bind specific molecular targets. Large random generated populations can be enriched in aptamers by *in vit* selection and polymerase chain reaction[1–11]. But so far sing stranded DNA has not been investigated for aptamer properti nor has a target protein been considered that does not inter physiologically with nucleic acid. Here we describe the isolati of single-stranded DNA aptamers to the protease thrombin of t blood coagulation cascade and report binding affinities in the ran 25–200 nM. Sequence data from 32 thrombin aptamers, select from a pool of DNA containing 60 nucleotides of random sequen displayed a highly conserved 14–17-base region. Several of the aptamers at nanomolar concentrations inhibited thrombin-ca lysed fibrin-clot formation *in vitro* using either purified fibrinog or human plasma.

We synthesized a pool of ~10^{13} 96-mer oligodeoxyribonucle tides that share 18-nucleotide binding sites for polymerase cha reaction (PCR) primers at their 5′ and 3′ termini and also conta 60-nucleotide randomly generated sequences. Using a 5′-biotir lated primer for one strand, this pool was amplified and rad labelled by PCR. The nonbiotinylated strand was isolated fr its complementary strand after application to an avidin–agarc column and base denaturation[12]. Single-stranded DNA was th applied to a concanavalin A (con A)–agarose column to remo DNA with affinity for con A, and the eluent applied to hum thrombin immobilized on con A. After washing unb und DN from the column, thrombin–DNA complexes were eluted w a con A ligand, α-methylmannoside. Fractions containi thrombin were identified using a chromogenic substrate ass and DNA from those fracti ns was quantitated bef re PC

* To whom correspondence should be addressed.

TABLE 1 Sequence specificity of thrombin inhibition

Sample	Sequence	Concentration	Clotting time (s) Purified fibrinogen	Human plasma
No DNA	—	—	25 ± 1	25 ± 1
Clone 29	96-mer	100 nM	76 ± 3	50 ± 2
Unselected DNA	96-mer	100 nM	26 ± 1	26 ± 1
Consensus 15-mer	GGTTGGTGTGGTTGG	100 nM	169 ± 8	43 ± 2
Scrambled 15-mer	GGTGGTGGTTGTGGT	100 nM	26 ± 1	26 ± 1
Consensus 6-mer	GGTTGG	20 μM	42 ± 2	40 ± 2
Scrambled 6-mer	TGGGGT	20 μM	26 ± 1	26 ± 1

DNA was incubated for 1 min at 37 °C in either selection buffer (0.2ml; Fig. 1 legend) containing human fibrinogen (Sigma) or fresh human plasma. Thrombin (0.1 ml in selection buffer pre-equilibrated to 37 °C) was added to give a final concentration of 13 nM thrombin and the indicated concentrations of DNA; the final concentration of purified fibrinogen was 2 mg ml^{-1}. Clotting times were measured using an automated fibrometer. Human blood was obtained by venipuncture, anticoagulated with the addition of 0.1 volumes 3.8% sodium citrate and fractionated by centrifugation (at 2,000*g* for 5 min). Plasma was decanted and stored (for less than one day) at 4 °C. Clotting times are the average of three experiments.

amplification and repeated selection. The scheme for the selection and amplification cycle is shown in Fig. 1.

Only 0.01% of the input DNA eluted with thrombin during the first selection cycle; this percentage increased in subsequent rounds of selection to ~40% by selection cycle 5. DNA from selection cycle 5 was assayed for thrombin binding specificity on nitrocellulose filters. The results showed that a significant fracti n (~30%) of aptamer DNA bound to thrombin but not to either ovalbumin or human fibrinogen (data not shown). Cycle-5 DNA was cloned in order to test the affinity of individual aptamers for thrombin, and radiolabelled single-stranded DNA from several clones was used to measure dissociation constants (K_d). Several clones gave $K_d \approx 200$ nM, whereas the original pool of DNA showed little affinity for thrombin (data not shown).

We determined the DNA sequence of the 60-nucleotide randomly generated region from 32 clones to examine both the heterogeneity of the selected population and to identify homologous sequences. Sequence analysis showed each of the 32 clones to be distinct, but there was a striking sequence conservation evident in every clone. The hexamer GGTTGG was found at a variable location within the 60-nucleotide randomized region in 31 out of 32 clones. The remaining clone contained the sequence TGTTGG, a match of five out of six bases. Additionally, in 28 of the 32 clones, a second hexamer GGNTGG was located 2–5 nucleotides 5′ or 3′ from the hexamer GGTTGG. The DNA sequence from the region of conservation from all 32 clones is presented in Fig. 2. Except for four clones that contain a close variation, the sequence $GGNTGGN_{2-5}GGNTGG$ is conserved. A compilation of the data relating base frequency to position is also shown in addition to the derived consensus sequence. DNA sequencing of clones (>20) from the unselected DNA population or from a population of aptamers selected for binding to a different target revealed no homology to the thrombin-selected aptamers (data not shown). These results indicate that this consensus sequence is responsible, either wholly or in part, for conferring thrombin affinity on the 96-mer oligodeoxyribonucleotide.

We next analysed whether the aptamer DNA could inhibit the thrombin-catalysed conversion of fibrinogen to fibrin using either purified human fibrinogen or adult human plasma. Using clone 29 DNA (Fig. 2) and purified fibrinogen, fibrin formation was detected at 76 seconds, compared with 26 s with unselected DNA or 25 s in the absence of DNA (Table 1). More impressively, with respect to oligonucleotide size and potency, fibrin formation using the 15-mer GGTTGGTGTGGTTGG (contained in clones 15 and 29) was detected at 169 s. The 6-mer GGTTGG also inhibited thrombin, although a higher concentration was required. No inhibition was seen with 15-mer or 6-mer scrambled sequence controls. The aptamers had anti-clotting activity in human plasma, although they were less potent (Table 1). This difference in potency may be attributable to nuclease

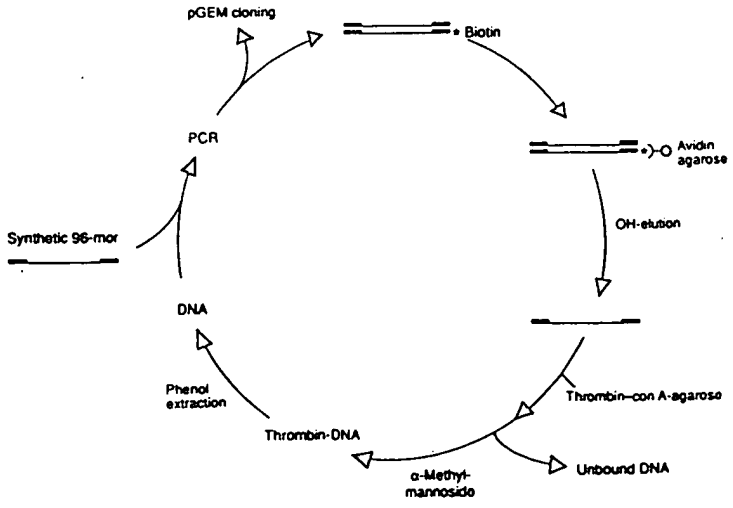


FIG. 1 Scheme for the selection of DNA aptamers to human thrombin.
METHODS. 96-mer DNA was prepared by solid-phase phosphoramidite chemistry on a Biosearch 8600 synthesizer using an equimolar mixture of the four bases for the 60-nucleotide random portion of the sequence which is flanked by defined 18-nucleotide regions that allow for PCR priming by the following oligonucleotides: 5′-CGTACGGTCGACGCTAGC-3′ and 5′ biotin-GGATCCGAGCTCCACGTG-3′ (biotinylation reagent from NEN). The synthetic DNA was purified by PAGE and amplified 100-fold by large-scale PCR to generate a sequence library with a complexity $>10^{13}$ individual sequences[1]. DNA from this library was then amplified and radiolabelled by PCR under standard conditions in the presence of 60 μCi [α-^{32}P]dNTPs. The resulting biotinylated (asterisk) double-stranded DNA was applied to an avidin–agarose (Vector Labs) column (>O) equilibrated with 0.1M Tris–HCl, pH 7.5, 0.1M NaCl. ssDNA was then eluted in 0.15N NaOH (ref. 12). The sample was neutralized with acetic acid, concentrated, and the DNA precipitated with ethanol. Following centrifugation, DNA was redissolved in selection buffer (20 mM Tris–acetate, pH 7.4, 140 mM NaCl, 5 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$) and 100 pmol applied to 1 ml concanavalin A–agarose (Vector Labs) equilibrated in selection buffer. The flow-through was applied to 6 nmol of human thrombin (Sigma) immobolized on 1 ml concanavalin A–agarose equilibrated in selection buffer. After washing with several column volumes of selection buffer, thrombin–DNA complexes were eluted by the addition of 0.1M α-methylmannoside to the selection buffer wash. Fractions containing thrombin were identified using a chromogenic pelleted by centrifugation, redissolved, and subjected to further cycles of amplification and selection. After the fifth selection cycle, DNA was amplified by PCR using the primers 5′-CGTACGGTCGACGCTAGC-3′ and 5′-TAATACGACTCACTATAGGGATCCGAGCTCCACGTG-3′, and subsequently cloned using the plasmid pGEM3Z (Promega). Single-stranded 96-mer DNA was then

Clone																		
1	g	G	G	t	T	G	G	-	-	g	t	c	G	G	t	T	G	G	t
2	g	G	G	a	T	G	G	-	-	t	t	t	G	G	t	T	G	G	g
3	a	G	G	t	T	G	G	-	-	-	g	a	G	G	g	T	G	G	g
4	t	G	G	t	T	G	G	-	-	c	g	a	G	G	a	T	G	G	a
5	a	G	G	t	T	G	G	-	g	t	a	g	t	G	t	T	G	G	t
6	a	G	G	t	T	G	G	-	-	g	c	t	G	G	t	T	G	G	g
7	g	G	G	t	T	G	G	-	-	-	g	a	G	G	t	T	G	G	a
8	t	G	G	t	T	G	G	-	-	g	t	c	G	G	t	T	G	G	g
9	g	G	G	a	T	G	G	-	-	t	g	t	G	G	t	T	G	G	c
10	t	G	G	t	T	G	G	-	-	c	a	g	G	G	a	T	G	G	g
11	t	G	G	a	T	G	G	-	-	t	g	a	G	G	t	T	G	G	a
12	g	G	G	g	T	G	G	-	-	t	t	a	G	G	t	T	G	G	t
13	a	G	G	g	T	G	G	-	-	t	t	a	G	G	t	T	G	G	t
14	c	G	G	t	T	G	G	-	g	t	t	g	G	G	a	T	G	G	a
15	c	G	G	t	T	G	G	-	-	t	g	t	G	G	t	T	G	G	t
16	a	G	G	t	T	G	G	-	-	t	g	t	G	G	g	T	G	G	g
17	c	G	G	g	T	G	G	-	-	a	t	a	G	G	t	T	G	G	a
18	g	G	t	g	T	G	G	t	a	g	t	t	t	G	t	T	G	G	g
19	t	G	G	t	T	G	G	t	t	a	c	t	G	G	t	T	G	G	g
20	g	G	G	t	T	G	G	-	-	t	c	t	G	G	g	T	G	G	a
21	t	G	G	t	T	G	G	-	-	g	t	t	G	G	g	T	G	G	a
22	t	G	G	t	T	G	G	-	-	c	c	a	G	G	t	T	G	G	a
23	c	t	a	g	c	G	G	-	c	a	g	t	G	G	t	T	G	G	g
24	t	G	G	g	T	G	G	-	-	g	g	a	G	G	t	T	G	G	t
25	a	G	G	t	T	G	G	-	-	t	t	t	G	G	g	T	G	G	t
26	a	G	G	t	T	G	G	-	t	t	a	g	G	G	t	T	G	G	t
27	g	G	G	a	T	G	c	-	-	g	g	t	G	G	t	T	G	G	g
28	t	G	G	t	T	G	G	-	t	t	a	t	G	G	t	T	G	G	t
29	a	G	G	t	T	G	G	-	-	t	g	t	G	G	t	T	G	G	c
30	a	G	G	t	T	G	G	-	-	t	g	t	G	G	g	T	G	G	g
31	t	G	G	t	T	G	G	-	-	-	g	a	G	G	t	T	G	G	t
32	g	G	G	t	T	G	G	t	g	g	g	t	G	G	a	T	G	G	t
G		31	30	6	0	32	31						30	32	6	0	32	32	
A		0	1	4	0	0	0						0	0	4	0	0	0	
T		1	1	22	31	0	0						2	0	22	32	0	0	
C		0	0	0	1	0	1						0	0	0	0	0	0	
Consensus sequence																			
		G	G	t	T	G	G			(N)3			G	G	t	T	G	G	

FIG. 2 Thrombin aptamer sequence homology. DNA from the fifth cycle of thrombin selection was amplified by PCR and cloned (see Fig. 1 legend). The sequences of the randomly generated 60-mer inserts were determined for 32 clones using dideoxynucleotide chain termination. A region of homology was identified shown in bold upper case. Below is a tabulation of the data and the deduced consensus sequence.

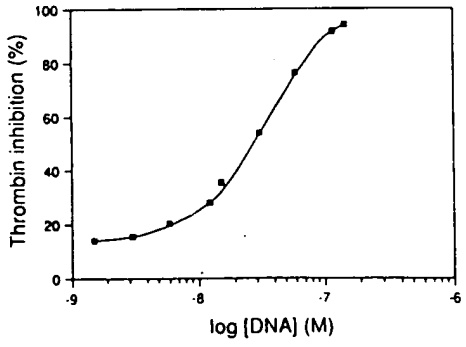


FIG. 3 Per cent thrombin inhibition versus DNA concentration.
METHODS. Selection buffer containing human fibrinogen (2mg ml^{-1} final) and varying concentrations of DNA (~1 nM–150 nM) was incubated for 1 min at 37 °C before adding thrombin as described in Table 1 legend. Clotting times were measured using an automated fibrometer. The extent of thrombin inhibition was then calculated using a thrombin standard curve generated by measuring clotting time versus thrombin concentration.

action or to binding of prothrombin or other plasma proteins present in much higher concentrations than thrombin. When the 6-mer was tested at much higher concentrations it was still as potent, perhaps because plasma nucleases and/or binding proteins were saturated. Thus, this protocol enabled us to isolate single-stranded DNA molecules that inhibit catalysis by thrombin. Moreover, this activity is retained by a short consensus sequence derived from the selected 96-mers.

We next determined the extent of thrombin inhibition with varying DNA concentration and found that thrombin was inhibited by 50% at 25 nM (Fig. 3). This is not a true inhibition constant, but it demonstrates the potency of the thrombin aptamer. Also, the data indicate that stoichiometry could be 1 : 1 and that each bound thrombin is largely, if not completely, inhibited.

Our findings show that a population of ~10^{13} molecules of 96-mer single-stranded DNA can be selected for aptamers that bind human thrombin, a protein with no known nucleic acid-binding function. As thrombin is a glycoprotein, we have been able to use lectin-agarose to immobilize it and α-methylmannoside to elute the thrombin-DNA complexes. Initially we bound DNA to thrombin that was covalently linked to agarose. Denaturing elution with EDTA gave single-stranded DNA with affinity for the matrix and hence only a modest enrichment of thrombin aptamers. These results suggest that the conditions under which aptamers are eluted from a covalently bound target are crucial to the successful isolation of high-affinity aptamers. One way to circumvent these problems may be to elute aptamer-target complexes from the matrix. Once isolated from matrix-associated oligonucleotides, the complexes can be fully denatured and the aptamers with the highest affinity recovered. In view of our success in isolating high-affinity aptamers to thrombin, we believe that lectin immobilization may provide an aptamer selection technique applicable to a large variety of glycoproteins.

The basic nature of thrombin may also have facilitated the isolation of aptamers; apart from the catalytic site, thrombin also contains an anion-binding exosite which binds fibrinogen[13–15], and a binding site for heparin, a polyanion. *In vitro* selection should allow the isolation of aptamers to most proteins, although the affinity may be highest for proteins rich in basic residues. The strong sequence dependence shown in Fig. 3 and in Table 1 suggests, however, that electrostatic interactions may be necessary but not sufficient for high-affinity thrombin binding.

It may be that the affinity of an aptamer for its ligand is comparable to that of an antibody for its antigen. The chemical nature, size and mode of isolation of aptamers may sometimes offer advantages over existing antibody technology. We are at present investigating the aptamer-binding site on thrombin and analysing the binding sequences of individual aptamers in an effort to understand the base relationships that mediate binding and inhibition, our long-term interest being to develop diagnostics and therapeutic agents. □

Received 28 August; accepted 15 November 1991.

1. Ellington, A. D. & Szostak, J. W. *Nature* **346**, 818–822 (1990).
2. Blackwell, T. K. & Weintraub, H. *Science* **250**, 1104–1110 (1990).
3. Blackwell, T. K. & Weintraub, H. *Science* **250**, 1149–1152 (1990).
4. Kinzler, K. W. & Vogelstein, B. *Nucleic Acids Res.* **17**, 3645–3653 (1989).
5. Thiesen, H. J. & Bach, C. *Nucleic Acids Res.* **18**, 3203–3208 (1990).
6. Kinzler, K. W. & Vogelstein, B. *Molec cell. Biol.* **10**, 634–642 (1990).
7. Chittenden, T., Livingston, D. M. & Kaelin, W. G. Jr *Cell* **65**, 1073–1082 (1991).
8. Joyce, G. F. *Gene* **82**, 83–87 (1989).
9. Robertson, D. L. & Joyce G. F. *Nature* **344**, 467–468 (1990).
10. Green, R., Ellington, A. D. & Szostak, J. W. *Nature* **347**, 467–468 (1990).
11. Tuerk, C. & Gold, L. *Science* **249**, 505–510 (1990).
12. Hultman, T., Stahl, S., Moks, T. & Uhlén, M. *Nucleosides and Nucleotides* **7**, 629–638 (1988)
13. Fenton, J. W. *Semin. Thromb. Hemostat.* **14**, 234–240 (1987).
14. Bode, W. *et al. EMBO J.* **8**, 3467–3475 (1989).
15. Rydel, T. J. *et al. Science* **249**, 277–280 (1990).

ACKNOWLEDGEMENTS. We thank H. Weintraub and D. Melton for review of the manuscript, V. Collins for administrative assistance, and all members of the Gilead team.

PATENT COOPERATION TREATY

DOCKETED FOR ATTORNEY ATTENTION

From the INTERNATIONAL SEARCHING AUTHORITY

To: KAREN L. ELBING
CLARK & ELBING, LLP
176 FEDERAL STREET
BOSTON, MASSACHUSETTS 02110

RECEIVED
JUL 21 2000
CLARK & ELBING

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

Applicant's or agent's file reference XXXXX-371WO2 00786/371W02	Date of Mailing (day/month/year) 19 JUL 2000
	FOR FURTHER ACTION See paragraphs 1 and 4 below
International application No. PCT/US00/04925	International filing date (day/month/year) 24 FEBRUARY 2000
Applicant THE GENERAL HOSPITAL CORPORATION	

1. [X] The applicant is hereby notified that the international search report has been established and is transmitted herewith.

Filing of amendments and statement under Article 19:
The applicant is entitled, if he so wishes, to amend the claims of the international application (see Rule 46):

When? The time limit for filing such amendments is normally 2 months from the date of transmittal of the international search report; however, for more details, see the notes on the accompanying sheet.

Where? Directly to the International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland
Facsimile No.: (41-22) 740.14.35

For more detailed instructions, see the notes on the accompanying sheet.

2. [ ] The applicant is hereby notified that no international search report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. [ ] **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

[ ] the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

[ ] no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in rules 90 *bis* 1 and 90 *bis* 3, respectively, before the completion of the technical preparations for international publication.

Within **19 months** from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within **20 months** from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer P. PONNALURI
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/220 (July 1998)★

(See notes on accompanying sheet)

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference XXXXX-371WO2	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
International application No. PCT/US00/04925	International filing date *(day/month/year)* 24 FEBRUARY 2000	(Earliest) Priority Date *(day/month/year)* 24 FEBRUARY 1999
Applicant THE GENERAL HOSPITAL CORPORATION		

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.
   [ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).
   b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing:
   [ ] contained in the international application in written form.
   [ ] filed together with the international application in computer readable form.
   [ ] furnished subsequently to this Authority in written form.
   [ ] furnished subsequently to this Authority in computer readable form.
   [ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
   [ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has b een furnished.
2. [ ] **Certain claims were found unsearchable** (See Box I).
3. [X] **Unity of invention is lacking** (See Box II).
4. With regard to the **title**,
   [X] the text is approved as submitted by the applicant.
   [ ] the text has been established by this Authority to read as follows:
5. With regard to the **abstract**,
   [ ] the text is approved as submitted by the applicant.
   [X] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.
6. The figure of the **drawings** to be published with the abstract is Figure No. 1
   [X] as suggested by the applicant.
   [ ] because the applicant failed to suggest a figure.
   [ ] because this figure better characterizes the invention.
   [ ] None of the figures.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4, 7-16, 19-23

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Box III TEXT OF THE ABSTRACT (Continuation of item 5 of the first sheet)

The abstract has been corrected to include figure 1 as requested by the applicants.

NEW ABSTRACT

The invention features a method of identifying a polypeptide which increases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter (as shown in Figure 1), whereby expression of the reporter gene is increased if the library includes a polypeptide which increases gene expression from the promoter; (b) determining whether the reporter gene expression is increased in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is increased, identifying a polypeptide of the library which increases reporter gene expression.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST 1.2, MEDLINE, SCISEARCH, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,874,304 A (ZOLOTUKHIN et al) 23 February 1999(23.02.99), see the entire document.	1-4, 7-16, 19-23
Y	CLEM, R.J et al. Prevention of Apoptosis by a Baculovirus Gene During Infection of Insect Cells. Science. 29 November 1991, Vol. 254, pages 1388-1390, especially page 1388.	1-4, 7-16, 19-23
Y	BIRNBAUM et al. An Apoptosis-inhibiting Gene from a Nuclear Polyhedrosis Virus Encoding a Polypeptide with a Cys/His Sequence Motifs. Journal of Virology. April 1994, Vol. 68, No. 4, pages 2521-2528, especially pages 2523-2524.	1-4, 7-16, 19-23

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 10 MAY 2000	Date of mailing of the international search report 19 JUL 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PASARELLI et al. A Baculovirus Gene Involved in Late Gene Expression Predicts a Large Polypeptide with a Conserved Motif of RNA Polymerases. Jounal of Virology. July 1994, vol. 68, No. 7, pages 4673-4678, see entire document.	1-4, 7-16, 19-23

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C12Q 1/00, 1/70, 1/66; C12N 7/00, 5/00, 5/02; G01N 33/554; A61K 38/00; C07H 21/02, 21/04, 21/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-4, 7-16, 19-23, drawn to a method for identifying a polypeptide which increases gene expression from a promoter.
Group II, claim(s) 5-16, 19-23, drawn to a method of identifying a polypeptide which modulates activation of a transcription factor activation domain.
Group III, claim(s) 17-23, drawn to a method of identifying a compound which modulates gene expression from a promoter.
Group IV, claim(s) 24, drawn to a method of determining whether a compound modulates NF-kB biological activity.
Group V, claim(s) 25-30, drawn to a method of determining a compound which modulates BCMA biological activity.
Group VI, claim(s) 31-32, drawn to a method of determining a compound NF-kB activity.
Group VII, claim(s) 33-37, drawn to a polypeptide which modulates NF-kB activity.
Group VIII, claim(s) 38-39, drawn to an NF-kB modulator.
Group IX, claim(s) 40-41, drawn to a method of activating NF-kB activity in a cell.
Group X, claim(s) 42-43, drawn to use of a BCMA polypeptide or nucleic acid.
The inventions listed as Groups I-X do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of group I is polypeptides which increase gene expression is not present in other groups, and also the polypeptides which increase gene expression from a promoter and method of identifying is known. Thus, the groups lack unity.

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: KAREN L. ELBING
CLARK & ELBING, LLP
176 FEDERAL STREET
BOSTON, MASSACHUSETTS 02110

PCT

NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing *(day/month/year)*	18 JUL 2001
Applicant's or agent's file reference 00786-371WO2	**IMPORTANT NOTIFICATION**

International application No.	International filing date *(day/month/year)*	Priority Date *(day/month/year)*
PCT/US00/04925	24 FEBRUARY 2000	24 FEBRUARY 1999

Applicant
THE GENERAL HOSPITAL CORPORATION

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI Telephone No. (703) 308-0196

TERRY J. DEY
PARALEGAL SPECIALIST
TECHNOLOGY CENTER 1600

Form PCT/IPEA/416 (July 1992) ★

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12Q 1/00, 1/70, 1/66, C12N 7/00, 5/00, 5/02, G01N 33/554, A61K 38/00, C07H 21/02, 21/04, 21/00

A1

(11) International Publication Number: WO 00/50633

(43) International Publication Date: 31 August 2000 (31.08.00)

(21) International Application Number: PCT/US00/04925

(22) International Filing Date: 24 February 2000 (24.02.00)

(30) Priority Data:
60/121,485 24 February 1999 (24.02.99) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 60/121,485 (CIP)
Filed on 24 February 1999 (24.02.99)

(71) Applicant *(for all designated States except US)*: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).

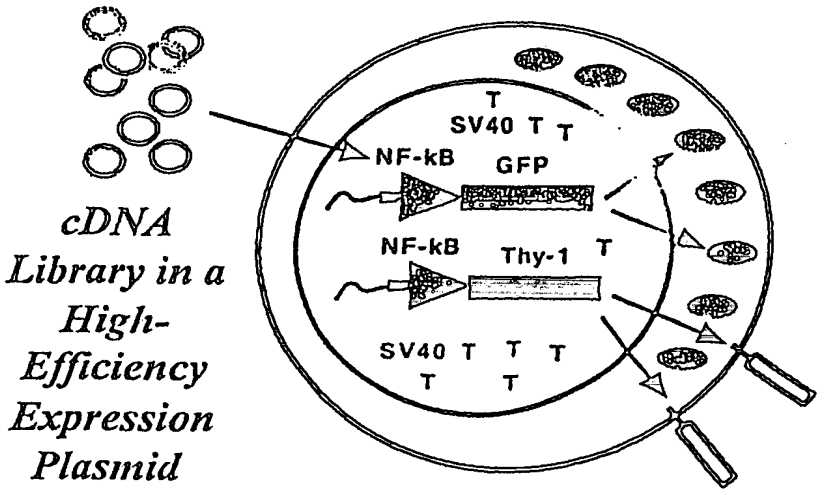
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SEED, Brian [US/US]; Apartment 5J, Nine Hawthorne Place, Boston, MA 02114 (US). TING, Adrian [MY/US]; 1675 York Avenue #6G, New York, NY 10128 (US).

(74) Agent: ELBING, Karen, L.; Clark & Elbing, LLP, 176 Federal Street, Boston, MA 02110 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD FOR CLONING SIGNAL TRANSDUCTION INTERMEDIATES



(57) Abstract

The invention features a method of identifying a polypeptide which increases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti–cell death gene and that contains a reporter gene operably linked to a promoter (as shown in the figure), whereby expression of the reporter gene is increased if the library includes a polypeptide which increases gene expression from the promoter; (b) determining whether the reporter gene expression is increased in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is increased, identifying a polypeptide of the library which increases reporter gene expression.

8/pats

METHOD FOR CLONING SIGNAL TRANSDUCTION INTERMEDIATES

Statement as to Federally Sponsored Research

This research has been sponsored in part by grant number AI27849 from the National Institutes of Health. The U.S. government has certain rights to this invention.

Background of the Invention

The invention relates to methods of identifying polypeptides and compounds which regulate gene expression.

Pharmaceuticals have historically been developed by testing libraries of up to several thousand compounds in laboratory animals, usually one compound at a time. The slow pace of the process, and its unsuitability for screening large numbers of diverse compounds, led to the development of approaches based on assays that can be completed quickly and *ex vivo*. With these approaches, the pharmaceutical drug discovery process has evolved into a catenation of several, sometimes partially concurrent, phases.

In the first phase, target identification and validation, a candidate target for a potential drug is identified by various means. These means may include hypotheses formed from the study of the pathophysiology of disease in humans or experimental animals, analysis of candidate signal transduction pathways *in vitro*, natural experiments such as genetic disorders of humans or other animals, results from targeted or random gene disruptions in model organisms, or disclosures by competitors. The projected consequences of a hypothetical drug interacting with its intended target to stimulate, block or

modulate the target activity is then tested. If the outcome of testing suggests that a drug affecting the target would have the desired physiological consequences in an intact organism, the target is said to be validated. Neither target identification nor validation are formally required, and, indeed, the historical process had neither of these phases.

Present models for pharmaceutical industry productivity emphasize the importance of efficiently identifying large numbers of validated targets and of developing rapid screens for their activity. The impetus for this model is the widespread belief that targets cannot be predicted effectively, so that a large number of targets must be evaluated to develop a small number of drugs. Hence, there remains an important need across the industry to accelerate the pace of target discovery and validation.

Once the target has been identified and validated, an assay must be created to allow large libraries of synthesized or natural compounds to be tested for their ability to interact with the target. When a compound is identified which acts on the target specifically, it is usually said to be a hit. The definition of a hit generally encompasses compounds that pass various secondary tests to assure that their activity is specific to the target of interest. From the various hits that are generated, one or more structures are chosen to represent the starting point for a program of systematic modification of the chemical structure. These structures are called lead compounds, and they are frequently selected from hits on the basis of their compatibility with directed synthesis programs, expected toxicity, or expected absorption, distribution, metabolism or excretion characteristics. Again, the process encompasses a wide range of practices and although they are usually distinguished, hits and lead compounds are sometimes referred to in ways that make them appear equivalent.

Summary of the Invention

In general, the invention provides a novel, rapid method for identifying members of selected signal transduction pathways which are targets for drug design.

In a first aspect, the invention features a method for identifying a polypeptide which increases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter, whereby expression of the reporter gene is increased if the library includes a polypeptide which increases gene expression from the promoter; (b) determining whether the reporter gene expression is increased in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is increased, identifying a polypeptide of the library which increases reporter gene expression.

In a second aspect, the invention features a method for identifying a polypeptide which decreases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to the promoter, whereby expression of the reporter gene is decreased if the library includes a polypeptide which decreases gene expression from the promoter;

(b) determining whether reporter gene expression is decreased in the cell as a result of contact with the polypeptide library; and

(c) if reporter gene expression is decreased, identifying a polypeptide which decreases reporter gene expression.

In a third aspect, the invention features a method for identifying a polypeptide which modulates activation of a transcription factor activation domain, including (a) contacting a library of polypeptides with a cell that

expresses a recombinant anti-cell death gene and a chimeric transcription factor consisting of a yeast or bacterial DNA binding domain fused to a mammalian transcription factor activation domain, and that contains a reporter gene operably linked to a promoter consisting of a basal promoter and binding sites for the DNA binding domain, whereby expression of the reporter gene is altered if the library comprises a polypeptide which modulates activation of the transcription factor activation domain; (b) determining whether reporter gene expression is altered in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is altered, identifying a polypeptide which modulates reporter gene expression.

In a fourth aspect, the invention features a method for identifying a compound which modulates gene expression from a promoter, including (a) contacting a library of compounds with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to the promoter, whereby expression of the reporter gene is altered if the library includes a compound which modulates gene expression from the promoter;

(b) determining whether reporter gene expression in the cell is altered as a result of contact with the compound library; and

(c) if reporter gene expression is altered, identifying a compound from the library which modulates reporter gene expression.

In a fifth aspect, the invention features a method for identifying a compound which decreases gene expression, including (a) contacting a library of compounds with a cell expressing (i) a recombinant anti-cell death gene; (ii) a second gene encoding a polypeptide; and (iii) a reporter gene that would have decreased expression if the function of the polypeptide was blocked; (b) determining whether expression of the reporter gene is decreased as a result of contact with the compound library; and (c) if expression of the reporter gene is

decreased, identifying a compound from the library which decreases reporter gene expression.

In preferred embodiments of the first, second, and third aspects, a library of DNA molecules encoding the library of polypeptides are expressed in a cell. The cell can be the same cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter (in which case the polypeptide is produced by the same cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter), or the polypeptide can be produced by a cell other than the cell that expresses a recombinant anti-cell death gene and that includes a reporter gene operably linked to a promoter; preferably the DNA molecules are expressed from a high-efficiency expression system.

In other preferred embodiments of the first, second, and third aspect, the library of DNA molecules is introduced to the cell by transfection, and the mean number of DNA molecules introduced by transfection to the cell is at least 25. Preferably, the mean number of DNA molecules introduced by transfection into the cell is at least 100 or even 500.

In still other preferred embodiments of the first, second, and third aspect, the polypeptide is selected from the group consisting of an extracellular ligand, a cell surface receptor, and a signal transduction intermediate, and the DNA molecules are expressed from a high-efficiency expression system.

In preferred embodiments of the first, second, third, fourth, and fifth aspects, step (c) includes (i) dividing the library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which activates reporter gene expression is identified; the promoter can derived from a mammal and the library of polypeptides can include polypeptides derived from a bacterium or a virus.

In other preferred embodiments of the first, second, third, fourth, and fifth aspects, the promoter is a heterologous promoter, the reporter gene is GFP, the anti-cell death gene is selected from the group consisting of bcl family members, IAP family members, and crmA, and the cell is selected from the group consisting of CHO, CD-1, Cos, 293, HeLa, BHK, or L cells.

In another aspect, the invention features a method for determining whether a compound modulates NF-kB biological activity, the method including the steps of: (a) providing a cell expressing a BCMA polypeptide; (b) contacting the cell with a candidate compound; and (c) measuring the level of expression of the BCMA polypeptide in the cell, wherein a change in the level of expression of the BCMA polypeptide in the cell, relative to a cell not contacted with the candidate compound, identifies the candidate compound as a compound that modulates NF-kB activity.

In still another aspect, the invention features a method for determining whether a compound modulates BCMA biological activity, the method including the steps of: (a) providing a BCMA polypeptide; (b) contacting the polypeptide with a candidate compound; and c) measuring the level of biological activity of the BCMA polypeptide, wherein a change in the level of biological activity of the BCMA polypeptide, relative to a polypeptide not contacted with the candidate compound, identifies the candidate compound as a compound that modulates BCMA activity. The BCMA polypeptide can be in a cell or in a cell-free system. A preferred BCMA biological activity is the modulation of NF-kB biological activity (e.g., the modulation of transcription by NF-kB) or NF-kB expression. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

In yet another aspect, the invention features a method for determining whether a compound modulates NF-kB activity, the method including the steps of: (a) providing a BCMA polypeptide; (b) contacting the polypeptide with a candidate compound; and (c) detecting the binding of the candidate compound to the polypeptide, wherein a candidate compound that binds to the polypeptide is a compound that modulates NF-kB biological activity. The BCMA polypeptide can be in a cell or in a cell-free system. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

In another aspect, the invention features a substantially purified polypeptide consisting of a BCMA polypeptide molecule lacking a BCMA extracellular domain. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2). Other preferred polypeptides are BCMA polypeptides that modulate NF-kB activity.

In yet another aspect, the invention features an NF-kB modulator including a BCMA polypeptide covalently linked to a heterologous compound. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2). Preferably, the modulator modulates NF-kB biological activity or expression.

The invention also features methods for activating NF-kB activity in a cell by contacting the cell with a recombinant BCMA polypeptide having NF-kB activating activity or by contacting the cell with a recombinant nucleic acid molecule encoding a BCMA polypeptide having NF-kB activating activity.

In related aspects, the invention features the use of BCMA polypeptides and nucleic acids for preparing pharmaceutical compositions for treating cancer or apoptosis.

By a "reporter gene" is meant a DNA or RNA sequence which encodes a reporter protein that is capable of being readily detected either inside or outside a cell.

The reporter gene is operably linked to a promoter which shows low spontaneous activation, so that the activity of the reporter protein in the presence of the activating polypeptide be at least two standard deviations greater than the activity in its absence. Various methods of increasing the sensitivity of reporter genes are known in the art, including: deletion from natural genes of inhibitory sequences, which may be found both within and external to the transcribed portion and which may affect mRNA formation, stability, or translational efficacy; addition of efficiently utilized introns to increase the rate of formation of mature mRNA; multimerization of upstream activation regions or binding sites for known transcriptional activators; increasing the copy number of the reporter gene; and protection of the activity of the reporter gene from adventitious stimulatory or inhibitory activity in flanking DNA by inserting the reporter gene between matrix attachment regions or chromatin insulator sequences.

Many different types of reporter proteins are known in the art. They frequently comprise proteins not normally found, or present in minor amounts, in some cells; they include enzymes that detoxify antimicrobial agents, such as aminoglycoside or aminocyclitol phosphotransferases or acetyltransferases, beta-lactamases or chloramphenicol acetyltransferase; enzymes of diverse origin that catalyze chromogenic, fluorogenic, or chemiluminescent reactions in the presence of exogenous substrates, such as beta-galactosidase, beta-

glucuronidase, alkaline phosphatase, catechol 2,3-dioxygenase, or various peroxidases; enzymes that catalyze photoreactions, such as bacterial or firefly luciferases; enzymes, like glycosyl transferases, that generate nonproteinaceous structures easily detected by antibodies, lectins, or cognate binding proteins; proteins easily monitored upon cell surface expression or secretion such as surface or secreted antigens for which corresponding antibodies or recognition proteins are known; and proteins which catalyze the synthesis of, or stoichiometrically embody, fluorescent structures, without exogenous substrates, such as the jellyfish fluorescent proteins (e.g., GFP).

By "operably linked" is meant that the gene for the reporter protein is positioned adjacent to a promoter which directs transcription of the gene and, ultimately, facilitates expression of the reporter protein.

By "promoter" is meant any minimal nucleic acid sequence sufficient to direct transcription of the reporter gene. The promoter is one which is activated by binding to a polypeptide. Examples of promoters useful in the invention are promoters which are normally linked to genes which are expressed when a cell is in a pathologic or disease state (e.g., cancer, inflammation, or due to bacterial or viral infection), and the protein products of which are directly or indirectly responsible for this state. Suitable promoters include, but are not limited to, the NF-kB promoter, the interleukin-2 promoter, and the HIV-1 long terminal repeat promoter.

A library of DNA molecules refers to a set of DNA molecules, each in a DNA expression vector. Preferably, the DNA expression vector displays high efficiency such that the level of expression is high. While in most cases, the library includes DNA molecules encoding tens, hundreds, or even thousands of different polypeptides, DNA molecules in a library can also encode only one polypeptide (for example, during the final steps of a sib

selection). A library with hundreds of different DNA molecules is considered to have greater "complexity" than a library with five different DNA molecules. Like DNA libraries, compound libraries can have different degrees of complexity. One feature of the invention is a method which allows for following an activity of interest through the sequential screening of libraries with less and less complexity.

By "candidate compound" is meant a chemical, be it naturally-occurring or artificial, that is surveyed for its ability to modulate BCMA or NF-kB biological activity. Candidate compounds may include, for example, peptides, polypeptides, antibodies (and fragments thereof), synthesized organic molecules, naturally occurring organic molecules, nucleic acid molecules, and components or derivatives thereof. The candidate compounds may be screened using any of the methods described herein using an animal, a cell, a lysate or extract derived from a cell, or a molecule derived from a cell. The measuring may be, for example, for the purpose of detecting altered gene expression, altered RNA stability, altered protein stability, altered protein levels, altered protein phosphorylation, or altered protein biological activity. The means for measuring may include, for example, antibody labeling, immunoprecipitation, phosphorylation assays, and methods known to those skilled in the art for detecting nucleic acids.

By "modulating" is meant changing, either by decrease or increase.

By "BCMA polypeptide" is meant a polypeptide that has substantial identity to human or mouse BCMA as shown in Figs. 7A and 7B, respectively, over a region of twenty consecutive amino acids and has a BCMA biological activity.

By "substantial identity" is meant that two polypeptide sequences, when optimally aligned, such as by the GAP or BESTFIT programs using

default gap weights, share at least 80% sequence identity, more preferably at least 90% sequence identity, and most preferably at least 95% or even 99% sequence identity or more.

By "heterologous compound" is meant a polypeptide, chemical, or detectable label (e.g., biotin) that is not naturally associated, in this case, with the polypeptide.

By "BCMA biological activity" is meant BCMA-mediated modulation of NF-kB biological activity, or expression or the binding of an antibody that specifically binds a BCMA polypeptide.

By "NF-kB biological activity" is meant any function performed by activated NF-kB.

The invention features a method of screening DNA and compound libraries for their ability to modulate reporter gene expression in a cell which is expressing a recombinant anti-cell death gene. The expression of this gene allows the cell to survive in conditions which would otherwise lead to its death and, as a result, failure of the screening procedure.

This method is suitable for identifying polypeptides that modulate transcription from a selected promoter. These polypeptides are, in turn, targets for drugs. The method is also suitable for identifying compounds which either mimic or block function of a polypeptide which itself modulates transcription from a selected promoter.

Hence, the invention features a new, efficient multistep method for identifying lead compounds which modulate expression from a promoter for development of pharmaceutical compounds.

The invention also features a new NF-kB activator, BCMA. As NF-kB is involved in numerous cellular processes and disease states, BCMA is useful for the treatment of disease. BCMA is also useful for the identification

of compounds that modulate its expression or biological activity. These compounds may be developed as drugs, or used as lead compounds for the purpose of identifying drugs.

Other features and advantages of the invention will be apparent from the following detailed description, the drawings, and the claims.

Brief Description of the Drawings

Fig. 1 is a schematic illustration of a reporter cell system for rapid identification of cDNAs encoding polypeptides of interest. A cDNA library is transfected into a cell which includes a reporter gene (e.g., GFP or Thy-1) operably linked to a promoter (in this case the NF-kB promoter). If the polypeptide expressed by one of the plasmids in the reporter cell is capable of inducing expression from the NF-kB promoter, the reporter gene is expressed, and the detectable marker protein is produced.

Fig. 2 is a schematic illustration of a cDNA library being divided into smaller libraries. Each of the libraries is then transfected into reporter cells. Libraries which induce reporter gene expression are then further divided until the cDNA(s) encoding the polypeptide which induces reporter gene expression are isolated.

Fig. 3 is a schematic illustration of three stages of enrichment for an intracellular signal transduction intermediate. The ratios represent the occurrence of the cDNA encoding the reporter gene-inducing polypeptide compared to the total cDNAs. Note that, during transfection, each cell receives more than one plasmid.

Fig. 4 is a schematic illustration showing how secreted factor can be distinguished from intracellular signal transduction intermediates such as receptors or kinases. Soluble ligands diffuse through the medium to engage

receptors on untransfected cells, which can lead to reporter gene expression in most, if not all, cells. Intracellular effectors are restricted to the transfected cells.

Fig. 5 is a schematic illustration of arrayed clones. N^2 elements, pooled into N columns and N rows, can be assayed with 2N transfections. In the example provided, 4 rows and 4 columns were positive. 4^2 or 16 transfections are then required to identify the intersections unambiguously.

Fig. 6 is a schematic illustration of an indexed library protocol. The method is similar to the one described in Fig. 1 to Fig. 5, except that individual plasmids are grown as individual cultures. Libraries of plasmids are then prepared and assayed. This method facilitates the rapid recovery of individual plasmids, and prevents the loss of positives when a library is divided into libraries of lower complexity.

Fig. 7A is a schematic illustration showing human BCNA polypeptide sequence (GenBank accession number Q02223). The putative transmembrane domain is indicated in bold.

Fig. 7B is a schematic illustration showing mouse BCNA polypeptide sequence (GenBank accession number AAC23799).

Fig. 8 is a schematic illustration showing that the intracellular domain of human BCMA was capable of activating NF-kB. A series of fusion proteins were constructed using the CD5L leader sequence, IgG Fc, a CD7 transmembrane domain, and amino acids from human BCMA (SEQ ID NO: 1) as follows: Ig7bcma--amino acids 78-184; ΔC20--amino acids 78-164; ΔC40--amino acids 78-144; ΔC60--amino acids 78-124; ΔC80--amino acids 78-104; ΔN20--amino acids 98-184; ΔN40--amino acids 118-184; ΔN60--amino acids 138-184; ΔN80--amino acids 158-184.

Detailed Description

The present invention provides a method for the rapid identification of molecules involved in regulation of specific signal transduction pathways. It can be used both as a target identification tool and as a rapid assay for drug effectiveness, and as such can significantly reduce the amount of time needed to go from target to hit.

The method features an expression cloning approach that identifies polypeptides that have the ability to activate reporter genes (Fig. 1). Such polypeptides are called here activating polypeptides. Expression cloning is a technique which identifies polypeptides solely on the basis of their ability to generate an observable activity of interest. In order for that activity to be identified, the polypeptide must generally be expressed, either *in vivo* or *in vitro*, and a suitably sensitive assay must be available to detect the activity after the polypeptide has been expressed. In the present method, it is preferred that the activity that is detected is dependent on the output of a reporter gene (Fig. 1).

Although the use of expression cloning in conjunction with reporter genes is not new, existing methods of application have not resulted in highly efficient systems for the identification of large numbers of new molecules. The present approach is a high throughput system for the identification of cDNA clones encoding polypeptides that induce reporter gene activity. Some features of the high throughput system of the present invention are: (1) a very high efficiency cDNA expression plasmid; (2) an easily detected reporter molecule; (3) features to prevent the death of the reporter cell due to toxicity of expressed genes; and (4) a method for the introduction of large numbers of plasmids into multiple cells.

Expression cloning can be carried out by either of two general

paradigms, positive selection, or sib selection, also known as pool division. By positive selection is meant a method for the enrichment of cells, viruses or genetically linked assemblies of proteins and the nucleic acids that encode them by physically separating the cells or genetically linked assemblies from the much larger collection of cells or assemblies that encode molecules of little interest. Because the nucleic acids that encode the polypeptides of interest are physically linked to the cells, viruses, or other assemblies that they comprise, the nucleic acid is physically purified by positive selection and the process can be repeated until a single nucleic acid is found that encodes the activity of interest. It is desirable in positive selection schemes to develop methods for association of a single nucleic acid with a single protein assembly or cell. Thus in positive selection it is desirable to identify and use methods for the introduction of nucleic acids into cells that result in a single nucleic acid species per cell.

The method of the invention uses sib selection, a process of screening in which pools of molecules that are derived from cognate libraries of nucleic acids are assayed for activity, and positive libraries are detected by some signal (e.g., expression of a reporter gene). The nucleic acids that make up the library are then separated into libraries with less complexity, which are then reassayed and redivided until a single nucleic acid is found that encodes the polypeptide with the desired activity (Figs. 2, 3, and 4) Because sib selection depends on the detection of the activity of aggregates of nucleic acids, it is often advantageous to use methods for the transfection or introduction of nucleic acids into cells that result in a large number of nucleic acid species per cell. Sib selection schemes, like positive selection methods, can also be carried out entirely *in vitro*.

Features of the Invention

Transfection Method

For sib selection to be an efficient method of cloning genes, according to the invention, it is important that a large number of DNA molecule are introduced into each cell. Many methods of introducing DNA molecules are known in the art, including microinjection, complexation with positively charged synthetic polycations such as DEAE dextran, polybrene, polylysine, or polyarginine, complexation with histones and other basic proteins, complexation with cationic lipids or related amphipathic molecules, condensation with polyethylene glycol or polyhydroxybutyrate, coprecipitation with calcium phosphate, electroporation, scrape loading or partial rupture, and fusion with bacterial or microbial spheroplasts. Preferred among these are methods that can be easily carried out in parallel and that result in the co-introduction into the reporter cell of multiple nucleic acids that encode different species. Moreover, the method of transfection will preferably provide, on average, at least 25 DNA molecules per cell. More preferably, the mean number will be at least 100 or even 500 DNA molecules per cell. The methods which achieve the preferred results include calcium phosphate coprecipitation, complexation with polycations or cationic lipids, and condensation with uncharged polymers such as polyethylene glycol.

The libraries of nucleic acid molecules encoding potentially activating polypeptides can also be created in a biologically active assembly. such as a virus or viral transducing particle, which is capable of introducing itself into the reporter cell directly. In such a case a cDNA library is prepared in the viral vector, and libraries of active virus or transducing particles are applied to the reporter cell. In general the reporter cell will have been previously engineered to contain a reporter gene, but the reporter gene may also

be introduced concurrently with the activating nucleic acid.

High Efficiency cDNA Expression Vector

To detect ectopic activation of a signal transduction pathway it is desirable to greatly overexpress proteins that act as signal transduction intermediates.

The invention uses a high efficiency cDNA expression system to produce proteins in the reporter cell. Such a system is provided by the use of a very strong promoter, such as, for example, the elongation factor 1 alpha (EF-1α) promoter; 3' untranslated region (3' UTR) and polyadenylation consensus (poly(A)) sequences from the human growth hormone gene; and the human IgG1 hinge-CH2 intron. Other strong promoter and nonpromoter elements are known in the art (for example, the murine or human cytomegalovirus immediate early gene promoters, globin introns, and 3' UTR/poly(A) sequences).

Cell Death Inhibition

Another important feature of the present system is a method to prevent the death of cells overexpressing proteins. Such a method is important for two reasons: signal transduction intermediates themselves can lead to cell death if they are expressed at high levels, and, in addition, if libraries of nucleic acids are transfected, the presence of even a low frequency of nucleic acids encoding toxic proteins can interfere with detection of the desired signal. The latter effect can be predicted to have greater impact as the size of the library increases.

There are several known methods to prevent the demise of cells undergoing programmed cell death, or apoptosis. Both viral and cellular

antagonists of cell death are known, and among them are species that act upstream in the apoptosis pathway, or at multiple levels. Examples of the former include dominant negative forms of cell death proteins, such as FADD or TRADD, and cellular inhibitors of apoptosis, such as members of the viral or cellular IAP family. Examples of proteins that act at multiple levels, or relatively downstream, include the orthopoxvirus CrmA and baculovirus p35 proteins, members of the cellular Bcl family, and both peptide and nonpeptide inhibitors of caspases, the cysteine proteases which act in a zymogen cascade to generate the death program. For the prevention of apoptosis in expression cloning, combinations of broad spectrum antiapoptotic proteins with different mechanisms of action are preferred, such as CrmA and Bcl-xL.

Examples

Identification of gene products leading to the activation of the transcription factor NF-kB

We established a reporter cell line by stably transfecting the commercially-available human embryonic kidney cell line, 293 EBNA, with two constructs: a reporter construct consisting of NF-kB promoter elements upstream of the green fluorescent protein (GFP); and a eukaryotic expression construct that expresses two anti-cell death molecules, CrmA and Bcl-xL. The former functions as a readout for the presence of signals that activate the NF-kB signal transduction pathway, whereas the latter prevents these cells from undergoing programmed cell death in the event that pro-apoptotic signals are present. The expression in the reporter cell line of NF-kB-activating molecules, including receptors (e.g. tumor necrosis factor-α receptor 1), ligands (e.g. tumor necrosis factor-α) and intracellular signaling proteins (e.g. RIP), results in the robust production of GFP. The level of GFP production, as a

measure of NF-kB activation, can be assayed using a fluorescent microscope or flow cytometry. High level expression of these various NF-kB-activating molecules was achieved by placing their cDNAs under the control of the very strong EF-1α promoter present in the commercially available plasmid PEAK8 (Edge Biosystems). Thus, one can transfect into these cells cDNA libraries and identify individual cDNAs which are capable of NF-kB activation.

A cDNA library, prepared from activated human T cell mRNA and cloned into the PEAK8 vector, was subdivided into smaller libraries of approximately 500 cDNA clones each, and DNA was prepared from each library. The DNA from the libraries were transfected into the reporter cell line by calcium phosphate precipitation following published protocols, e.g., Ausubel et al. et al., 1997, *Current Protocols in Molecular Biology*, Wiley Interscience. Cells were allowed to continue to grow for 48 hours and assayed for GFP production. cDNA libraries that gave a positive GFP signal, as defined by the appearance of at least 0.1% of bright cells by fluorescence microscopy, were screened further by sub-dividing into libraries of 50 cDNAs each (Fig. 2). DNA from each library was prepared, transfected into the reporter cell line and assayed for GFP production. The process was repeated until a single cDNA clone that induced a positive GFP readout was obtained. The screening procedure has resulted in the cloning of DNAs encoding known NF-kB-activating molecules such as surface receptors (DR3, FAS), soluble ligands (interleukin-1, TRAIL, CD40 ligand), intracellular signaling molecules (small molecular weight GTPase rho). Also identified were a previously known molecule having no known function (BCMA) and novel molecules.

The inclusion of two anti-cell death genes in the transfected cells are likely to have aided in the expression cloning of at least some of the above-mentioned DNAs. It has been established that expression of either FAS or DR3

in a cell would induce apoptosis in that cell. The fact that each was retrieved substantiates the improved nature of the expression cloning method of the present invention. Similarly, as the method of the present invention preferably employs high copy number, it is likely that expression of anti-cell death genes allows for cells that otherwise would have died, due to a high number of inserts, to survive.

The method is more rapidly and efficiently applied by using indexed arrays of bacterial cultures, in which each culture is derived from a single colony, and, thus, represents an independently derived cDNA expression plasmid (Fig. 5). By pooling small amounts of the cultures from rows and columns of large arrays and preparing DNA from each of the libraries, it is possible to assess the relative activity of every row and every column, thereby reducing labor dramatically (Fig. 6). For example if there are 90,000 individual/clones arrayed in a 300 by 300 matrix, the 300 rows and 300 columns can be transfected to determine all of the rows and all of the columns that bear positive clones. Hence the array can be indexed in only 600 DNA preparations and transfections. If there is only one positive culture, the job is complete. If there are multiple positive cultures, then in the worst case, each row and each column will have only one positive culture, and if there are x positives, then there are x^2 possible intersections. As long as x is a small number, though, the work involved is quite modest. For example if x were 10, the entire array of 90,000 clones could be screened in only 700 DNA preparations and transfections.

Variations of the experimental approach outlined above are applicable depending on the particular system that is to be examined. The reporter system and cell lines can be adapted to the promoter to be investigated. For instance, the NF-kB promoter in the GFP reporter construct can be replaced

by the interleukin-2 promoter, introduced into a T lymphocyte cell line and screened for molecules that regulate the signal transduction pathway leading to the expression of interleukin-2 in T lymphocytes. Other promoters can be drawn from viral sources, such as the HIV-1 long terminal repeat (LTR) promoter, or other inducible promoters of viral origin. The calcium phosphate transfection procedure can be changed to another method more suitable for the particular reporter cell type. The alternative transfection protocols could be based on electroporation, cationic lipids, DEAE-dextran, spheroblasts fusion or viral-mediated delivery.

A variation to the method described above is to screen for molecules that will turn off expression from a promoter. For instance, the NF-kB reporter in the reporter cell line is activated in response to interleukin-1 stimulation. By including interleukin-1 in the culture medium, one can then search for molecules that will inhibit the interleukin-1 dependent expression from the NF-kB promoter. Depending on the cDNA library that is used, these molecules can be either ones that are naturally negative regulatory or they can be mutant versions that behave in a dominant inhibitory manner. Similarly, screens can be conducted with a combinatorial library to look for small pharmacological molecules that will negatively interfere with the pathway (see below).

The screening methods described above are well-suited for screening for genes from one organism that interact with a pathway in another organism. A prime example is screening the expressed genomes of viruses, bacteria or other pathogens for genes that, when expressed, might interact with the NF-kB pathway, This method, utilized by these pathogens to alter the immune response to their advantage, can identify potential targets for pharmacological interventions.

The reporter cell may also be provided with proteins which increase

the efficiency of the reporter gene. These proteins are usually introduced by transfection of an expression plasmid that encodes them. Proteins which increase the efficiency of the reporter gene may provide essential elements of a signaling pathway which are not otherwise present in the reporter cell, or may consist of artificial proteins that act to integrate, amplify, or selectively respond to signals from the pathway of interest. Of special relevance are artificial transcriptional activators which consist of a DNA binding element that interacts specifically with cognate binding sites in an artificial promoter and a pathway-specific transcriptional activator element that responds to activation of the pathway with a change in transcriptional activity. Such artificial activators are known in the art as fusions between bacterial or yeast DNA binding proteins and mammalian transcription factor activation domains. Such activation domains provides sites for protein binding, ligand-activated conformational change or post-translational modification, that increase the transcription-promoting capabilities of the artificial activator. Examples of artificial activators that are commercially available or known in the art include LexA and Gal4 fusions with c-Jun, Elk1, CREB, c-Fos, ATF2, CHOP, and members of the nuclear hormone superfamily.

In one example, a cell which contains (i) a recombinant anti-cell death gene, (ii) a chimeric transcription factor consisting of the Gal4 DNA binding domain fused to the c-Jun activation domain, and (iii) DNA encoding GFP operably linked to regulatory sequence consisting of a basal promoter and Gal4 binding sites is used to identify polypeptides or compounds which modulate c-Jun activation. If, for example, expression of a polypeptide leads to c-Jun phosphorylation (i.e., activation), then increased GFP expression would result.

BCMA is an activator of NF-kB

Using the methods described herein, we identified BCMA as an NF-kB activator. BCMA was discovered by molecular analysis of a t(4;16) translocation characteristic of a human T cell lymphoma (Laabi et al., EMBO J. 11: 3897-3904, 1992), and its function was not previously known.

Based on the present discovery, BCMA is a target for drug discovery or rational drug design. For example, a compound that modulates BCMA expression or biological activity will also modulate NF-kB biological activity. Accordingly, the invention features methods and reagents for the identification of NF-kB modulating compounds.

BCMA polypeptides or nucleic acid molecules are also useful for the treatment of diseases associated with insufficient or inappropriate NF-kB biological activity or expression. BCMA polypeptides or nucleic acid molecules are administered to a patient using an appropriate delivery vehicle, as known in the art. Generally, the BCMA polypeptide or nucleic acid molecule is delivered in a pharmaceutically acceptable carrier.

Coupled Target Identification and Assay Generation

Directly relevant to the NF-kB signal transduction pathway is the identification of molecules that could potentially play a role in regulating inflammation and oncogenesis. The search for NF-kB activating molecules do not have to be restricted to screening cDNA expression libraries; the same principles can be used to identify compounds which modulate the output of the pathway, either by mimicking the activity of a polypeptide or by blocking its activity when the polypeptide is overexpressed. The latter compounds will act either upon or downstream of the overexpressed protein; once enough activating proteins have been identified, it will be possible to identify at what

step the compounds act by evaluating their action on a panel of transfected cells expressing different activating proteins.

In general, compounds are identified from large libraries of both natural product and synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceangraphics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (e.g., taxonomic

dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known to modulate the test promoter should be employed whenever possible.

When a crude extract is found to modulate reporter gene expression, further fractionation of the positive lead extract is necessary to isolate chemical constituents responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having the desired activity. Methods of fractionation and purification of such heterogenous extracts are known in the art. If desired, compounds shown to be useful agents for the modulation of gene expression or biological activity are chemically modified according to methods known in the art.

Primary screens for compounds that modulate BCMA biological activity

Modulating NF-kB expression or biological activity modulates numerous important cellular processes. The finding that NF-kB activity is regulated by BCMA allows us to provide assays for drugs that modulate NF-kB by monitoring BCMA expression or biological activity. Such assays may measure BCMA expression by measuring changes in: (a) levels of BCMA protein; (b) levels of BCMA RNA; (c) levels of BCMA-mediated NF-kB biological activity; or (d) levels of a reporter gene or protein expressed from a NF-kB promoter. These measurements may be made *in vitro* or *in vivo*. These assays allow for the identification of compounds that modulate NF-kB biological activity (e.g., gene transcription). Such identified compounds may have therapeutic value, for example, in the treatment of diseases that result in too little or too much cell death.

Cells overexpressing BCMA can be produced using standard techniques. Compounds that are identified may bind to BCMA and prevent BCMA activation of NF-kB. While screening of compounds can be performed with cultures of primary cells, cell lines be also be used. Cell lines can be modified such that the cells constitutively express a BCMA polypeptide, for example, the BCMA intracellular domain.

Any cell line, such as ones described above, can also be engineered to contain a reporter gene expressed under control of the NF-kB promoter (described above). A preferred reporter gene codes for GFP. Typically, the expression of the gene (e.g., the endogenous NF-kB gene or a recombinant reporter gene expressed under the control of the NF-kB promoter or fragment thereof) is measured by assaying the RNA or protein levels or both of the expressed gene. For example, the polypeptide expressed by the NF-kB gene or by the reporter gene produces a detectable signal under conditions that allow compound-mediated changes to be measured. Quantitatively determining the amount of signal requires comparing the amount of signal produced in the absence of any compound being tested to the amount produced when the cell is contacted with the compound, as is described herein. The comparison permits the identification of the compound as one that causes a change in the detectable signal produced by the expressed gene (e.g., at the RNA or protein level) and thus identifies a compound that is capable of modulating NF-kB expression. In order to prevent the NF-kB cells from dying, a second gene encoding an apoptosis inhibitor can also be expressed in the cells, as described herein.

Secondary screens for compounds that modulate NF-kB activity

After test compounds that appear to modulate NF-kB expression are identified, it may be necessary or desirable to subject these compounds to

further testing. The invention provides such secondary confirmatory assays. For example, a compound that appears to modulate NF-kB activity in early testing may be subject to additional assays to confirm that the compound also modulates NF-kB activity *in vivo*. In the first round of *in vivo* testing, NF-kB activity is initiated in animals by well-known methods and the compound is administered by one of the means described in the "Therapy" section, immediately below. Cells or cellular tissue are isolated within hours to days following the insult, and are subjected to assays to assess the level of NF-kB expression or biological activity. Such assays are well known to those skilled in the art. Examples of such assays include, but are not limited to, ELISAs, Western blot analysis, RT-PCR, RIA, and Northern blot analysis.

Therapy

NF-kB is an important regulator of inflammatory responses (e.g., rheumatoid arthritis, inflammatory bowel disease, septic shock), apoptosis, oncogenesis, and anti-viral and anti-bacterial responses. Therefore, the discovery of new gene products that regulate NF-kB activity, and thus the disease process, will result in the identification of molecular targets for pharmacological intervention. By increasing or mimicking BCMA biological activity, one could, for example, boost anti-tumor antibody production or increase T cell cytotoxicity against tumor cells in cancer immunotherapy. Conversely, antagonizing BCMA biological activity would be advantageous, for example, in situations in which it is desirable to down-regulate immune cell function. Compounds, identified using any of the methods disclosed herein, may be administered to patients or experimental animals with a pharmaceutically- acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable

formulations or compositions to administer such compositions to patients or experimental animals. Although intravenous administration is preferred, any appropriate route of administration may be employed, for example, parenteral, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Methods well known in the art for making formulations are found in, for example, Remington: The Science and Practice of Pharmacy, (19th ed.) ed. Gennaro AR., 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethyl ene-polyoxypropyl ene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for antagonists or agonists of the invention include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

As described herein, we have discovered that BCMA activates NF-kB activity. NF-kB, in turn, activates numerous cellular processes. Hence, any compound that modulates NF-kB expression is a candidate compound for use

in drug development. One possible compound is a polypeptide fragment of BCMA which maintains its ability to bind to another component of the signal transduction pathway but has lost its ability to activate NF-kB. Such a polypeptide will act as an inhibitor of wild-type NF-kB signaling . Another possible compound is a polypeptide fragment of BCMA which exhibits constitutive activation of NF-kB.

Other Embodiments

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure come within known or customary practice within the art to which the invention pertains and may be applied to the essential features hereinbefore set forth.

What is claimed is:

1. A method for identifying a polypeptide which increases gene expression from a promoter, said method comprising the steps:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is increased if said library comprises a polypeptide which increases gene expression from said promoter;

(b) determining whether said reporter gene expression is increased in said cell as a result of contact with said polypeptide library; and

(c) if said reporter gene expression is increased, identifying a polypeptide of said library which increases said reporter gene expression.

2. The method of claim 1, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which activates reporter gene expression is identified.

3. A method for identifying a polypeptide which decreases gene expression from a promoter, said method comprising:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is decreased if said library comprises a polypeptide which decreases gene expression from said promoter;

(b) determining whether said reporter gene expression is decreased in said cell as a result of contact with said polypeptide library; and

(c) if reporter gene expression is decreased, identifying a polypeptide

which decreases said reporter gene expression.

4. The method of claim 3, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which decreases reporter gene expression is identified.

5. A method for identifying a polypeptide which modulates activation of a transcription factor activation domain, said method comprising:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and a chimeric transcription factor consisting of a yeast or bacterial DNA binding domain fused to a mammalian transcription factor activation domain, and that comprises a reporter gene operably linked to a promoter consisting of a basal promoter and binding sites for said DNA binding domain, whereby expression of said reporter gene is altered if said library comprises a polypeptide which modulates activation of said transcription factor activation domain;

(b) determining whether said reporter gene expression is altered in said cell as a result of contact with said polypeptide library; and

(c) if reporter gene expression is altered, identifying a polypeptide which modulates said reporter gene expression.

6. The method of claim 5, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which modulates reporter gene expression is identified.

7. The method of claim 1, 3, or 5, wherein said contacting comprises expressing a library of DNA molecules in a cell, wherein said library of DNA molecules encodes said library of polypeptides.

8. The method of claim 7, wherein said cell is the same cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

9. The method of claim 7, wherein said library of DNA molecules is introduced to said cell by transfection.

10. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 25.

11. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 100.

12. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 500.

13. The method of claim 7, wherein said DNA molecules are expressed from a high-efficiency expression system.

14. The method of claim 1, 3, or 5, wherein said polypeptide is produced by the same cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

15. The method of claim 1, 3, or 5, wherein said polypeptide is produced by a cell other than the cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

16. The method of claim 1, 3, or 5, wherein said polypeptide is selected from the group consisting of an extracellular ligand, a cell surface receptor, and a signal transduction intermediate.

17. A method for identifying a compound which modulates gene expression from a promoter, said method comprising:

(a) contacting a library of compounds with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is altered if said library comprises a compound which modulates gene expression from said promoter;

(b) determining whether said reporter gene expression in said cell is altered as a result of contact with said compound library; and

(c) if said reporter gene expression is altered, identifying a compound from said library which modulates said reporter gene expression.

18. The method of claim 17, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a compound which modulates gene expression from a promoter is identified.

19. The method of claim 1, 3, 5, or 17, wherein said promoter is a heterologous promoter.

20. The method of claim 1, 3, 5, or 17, wherein said promoter is derived from a mammal and said library of polypeptides comprises polypeptides derived from a bacterium or a virus.

21. The method of claim 1, 3, 5, or 17, wherein said reporter gene is GFP.

22. The method of claim 1, 3, 5, or 17, wherein said anti-cell death gene is selected from the group consisting of bcl family members, IAP family members, and crmA.

23. The method of claim 1, 3, 5, or 17, wherein said cell is selected from the group consisting of CHO, CD-1, Cos, 293, HeLa, BHK, or L cells.

24. A method for determining whether a compound modulates NF-kB biological activity, said method comprising the steps of:

a) providing a cell expressing a BCMA polypeptide;

b) contacting said cell with a candidate compound; and

c) measuring the level of expression of said BCMA polypeptide in said cell, wherein a change in the level of expression of said BCMA polypeptide in said cell, relative to a cell not contacted with said candidate compound, identifies said candidate compound as a compound that modulates NF-kB biological activity.

25. A method for determining whether a compound modulates BCMA biological activity, said method comprising the steps of:

a) providing a BCMA polypeptide;

b) contacting said polypeptide with a candidate compound; and

c) measuring the level of biological activity of said BCMA polypeptide, wherein a change in the level of biological activity of said BCMA polypeptide, relative to a polypeptide not contacted with said candidate compound, identifies said candidate compound as a compound that modulates BCMA biological activity.

26. The method of claim 25, wherein said BCMA polypeptide is in a cell.

27. The method of claim 25, wherein said BCMA polypeptide is in a cell-free system.

28. The method of claim 25, wherein said BCMA biological activity is the modulation of NF-kB biological activity.

29. The method of claim 28, wherein said NF-kB biological activity is the modulation of cell death.

30. The method of claim 25, wherein said BCMA polypeptide comprises a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

31. A method for determining whether a compound modulates NF-kB activity, said method comprising the steps of:

a) providing a BCMA polypeptide;

b) contacting said polypeptide with a candidate compound; and

c) detecting the binding of said candidate compound to said polypeptide, wherein a candidate compound that binds to said polypeptide is a compound that modulates NF-kB biological activity.

32. The method of claim 31, wherein said BCMA polypeptide comprises a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

33. ~~32.~~ A substantially purified polypeptide comprising a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) and not having amino acids 1-54 of human BCMA (SEQ ID NO: 1).

34 ~~33.~~ The polypeptide of claim 32, wherein said polypeptide modulates NF-kB activity.

35. ~~34.~~ The polypeptide of claim 32, wherein said polypeptide consists of amino acids 98-164 of human BCMA (SEQ ID NO: 1).

36 ~~35.~~ A substantially purified polypeptide comprising a polypeptide sequence having substantial identity to amino acids 97-163 of mouse BCMA (SEQ ID NO: 2) and not having amino acids 1-49 of mouse BCMA (SEQ ID NO: 2).

37 ~~36.~~ The polypeptide of claim 35, wherein said polypeptide

modulates NF-kB activity.

38. The polypeptide of claim 35, wherein said polypeptide consists of amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

39. An NF-kB modulator comprising a polypeptide of claim 32 or 35 covalently linked to a heterologous compound.

40. The modulator of claim 38, wherein said modulator modulates NF-kB activity.

41. A method for activating NF-kB activity in a cell, comprising contacting said cell with a recombinant BCMA polypeptide having NF-kB activating activity.

42. A method for activating NF-kB activity in a cell, comprising contacting said cell with a recombinant nucleic acid molecule encoding a BCMA polypeptide having NF-kB activating activity.

43. Use of a BCMA polypeptide for preparing a pharmaceutical composition for treating cancer, apoptosis, a viral infection, or an inflammatory response.

44. Use of a BCMA nucleic acid molecule for preparing a pharmaceutical composition for treating cancer, apoptosis, a viral infection, or an inflammatory response.

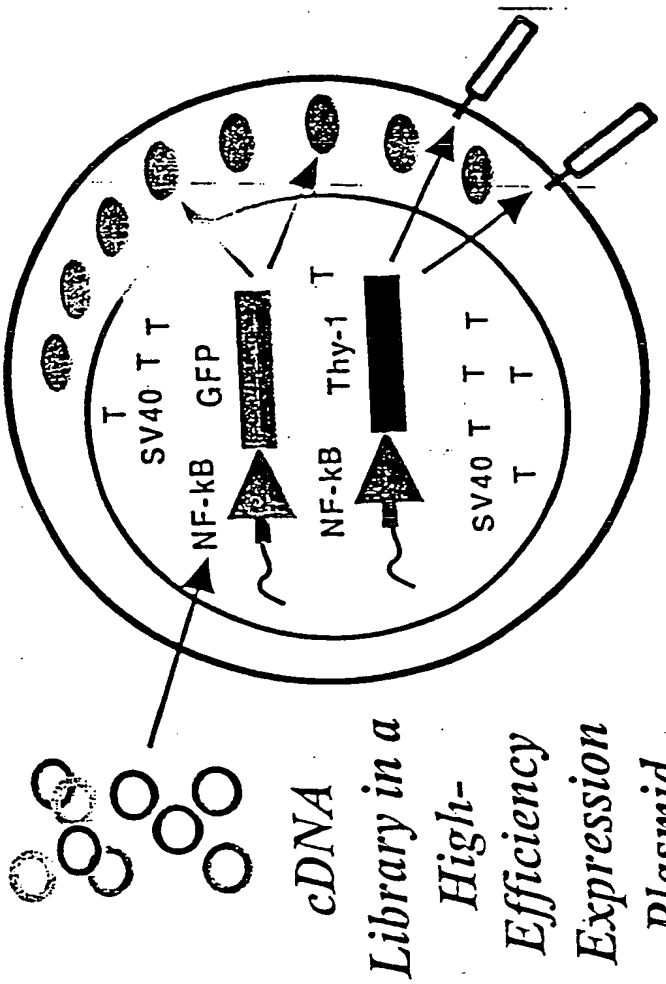
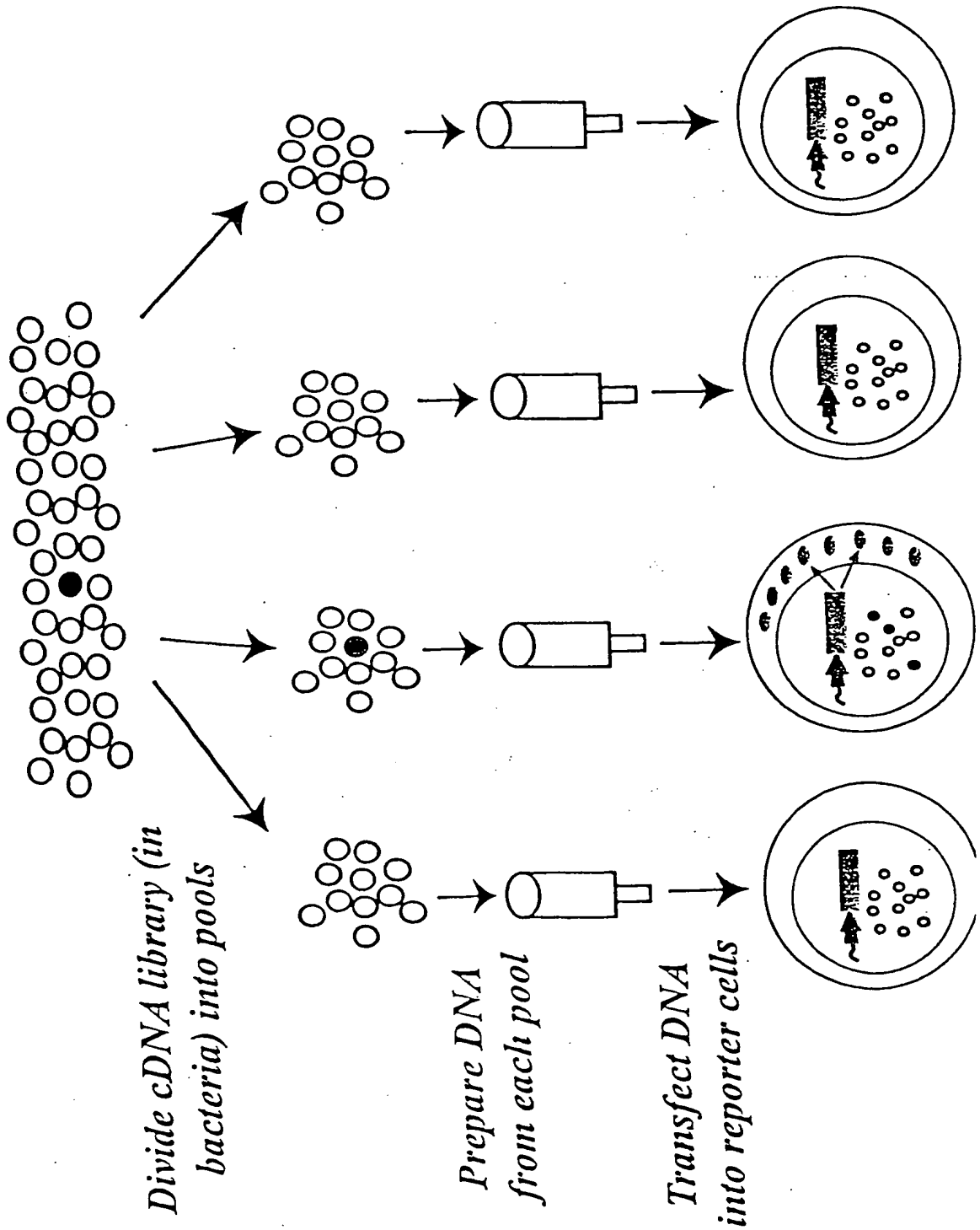

T
SV40 T T
GFP
NF-kB
T
Thy-1
NF-kB
SV40 T T T
T
cDNA
Library in a
High-
Efficiency
Expression
Plasmid

Fig. 1

Fig. 2


Divide cDNA library (in bacteria) into pools
Prepare DNA from each pool
Transfect DNA into reporter cells

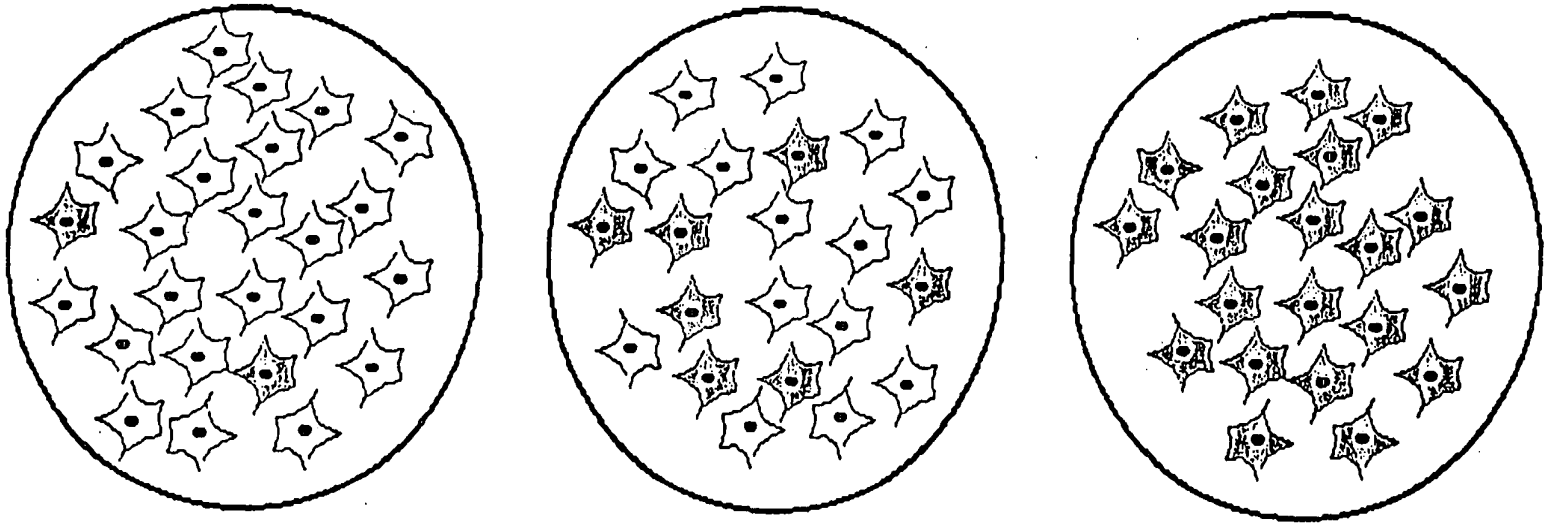


Fig. 3

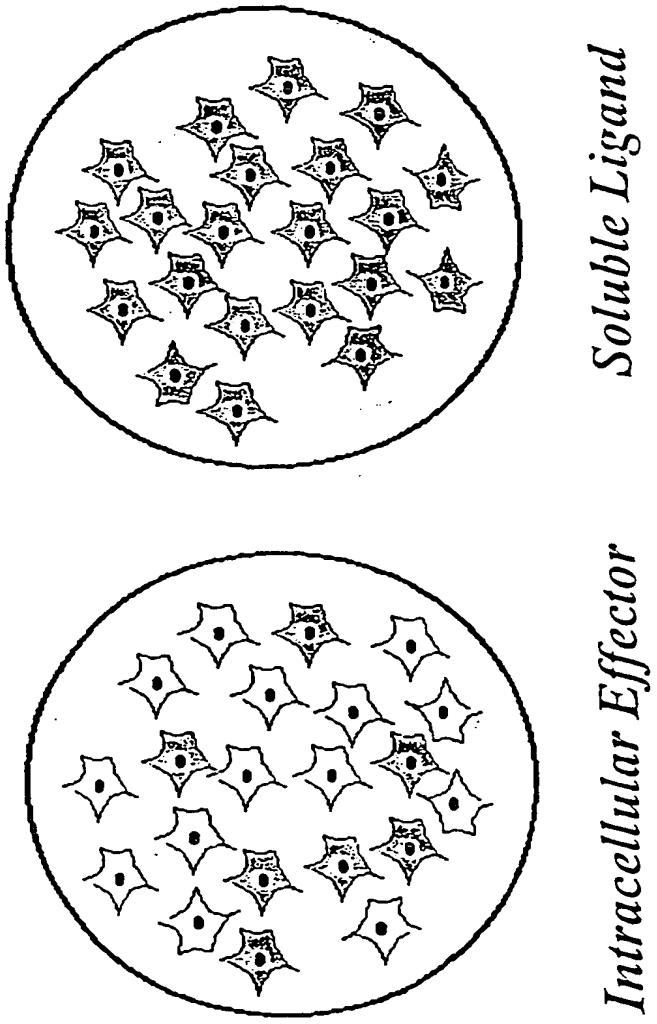


Fig. 4

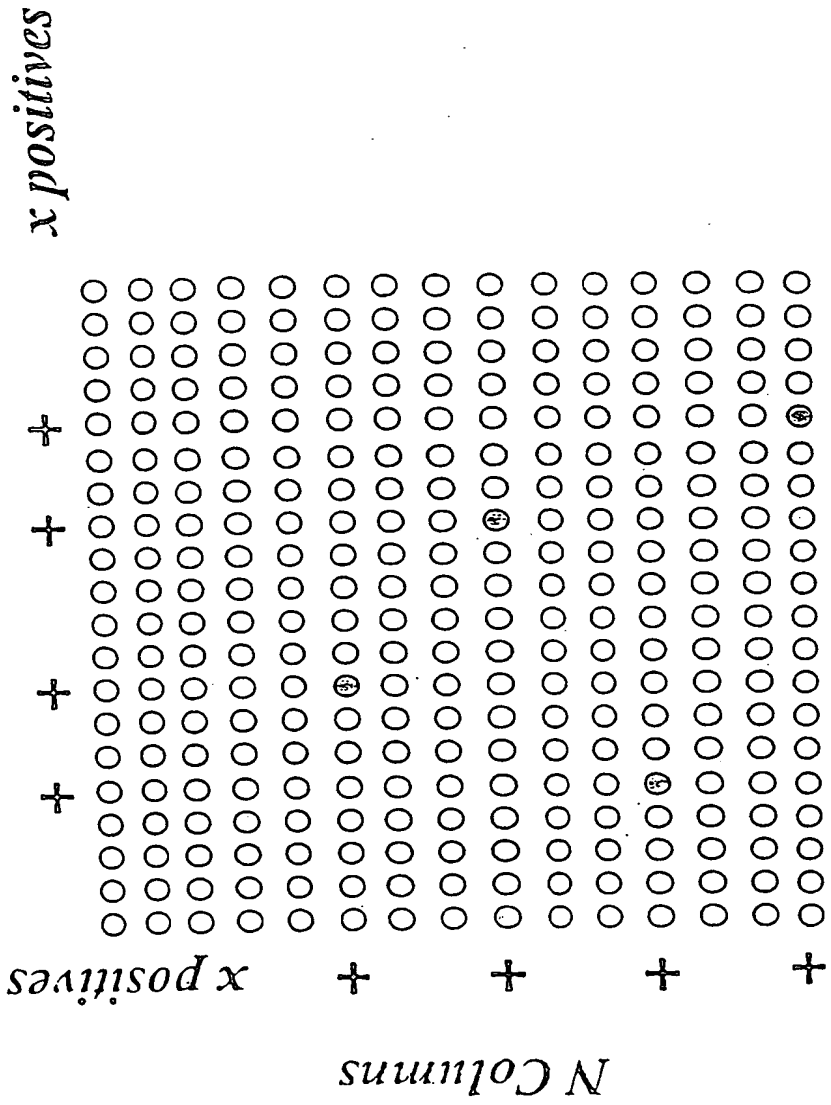


Fig. 5

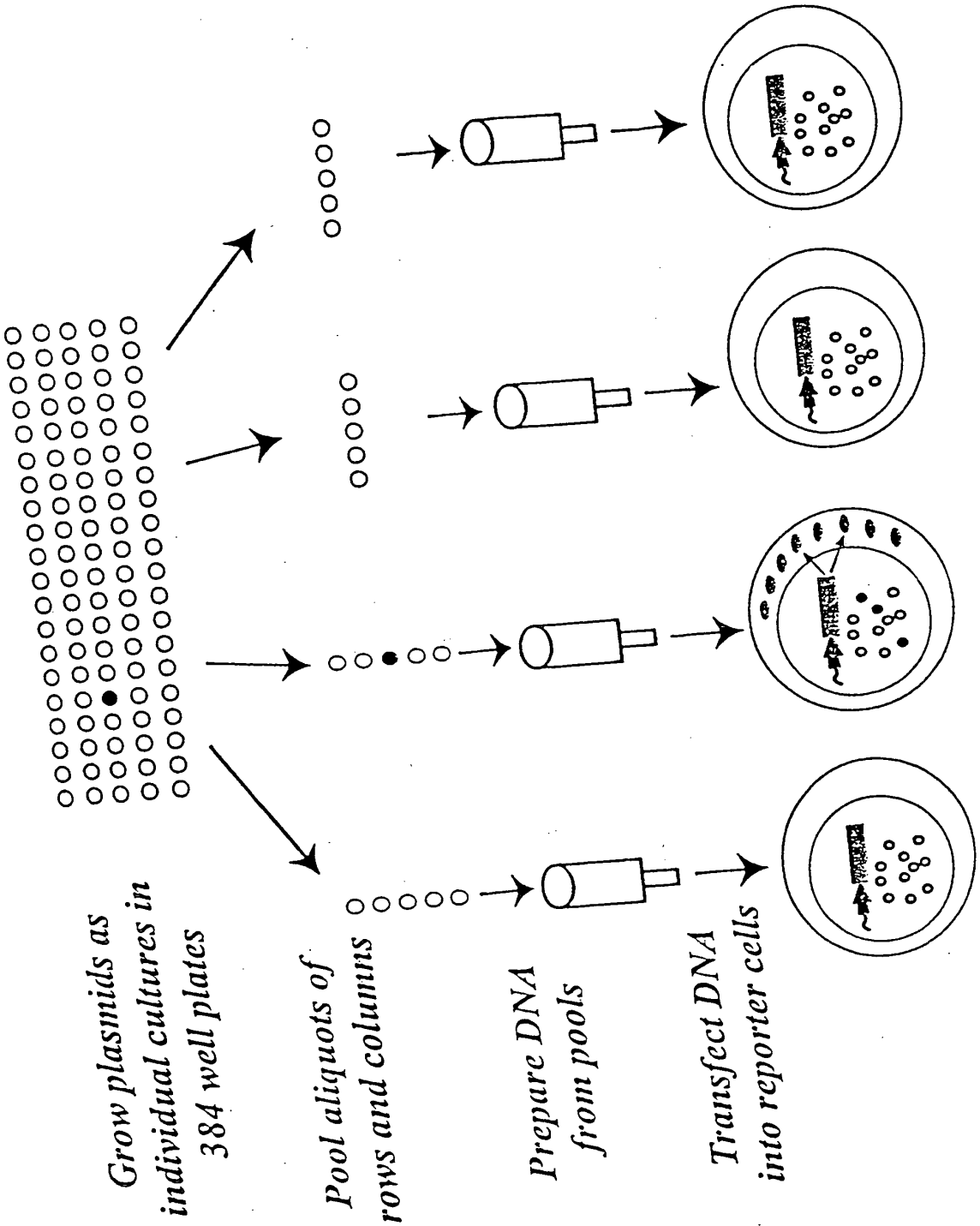

Grow plasmids as individual cultures in 384 well plates
Pool aliquots of rows and columns
Prepare DNA from pools
Transfect DNA into reporter cells

Fig. 6

MLQMAGQCSQ NEYFDSLLHA CIPCQLRCSS NTPPLTCQRY CNASVTNSVK GTNAILWTCL
GLSLIISLAV FVLMFLLRKI SSEPLKDEFK NTGSGLLGMA NIDLEKSRTG DEIILPRGLE
YTVEECTCED CIKSKPKVDS DHCFPLPAME EGATILVTTK TNDYCKSLPA ALSATEIEKS
ISAR (SEQ ID NO: 1)

Fig. 7A
Human BCNA

MAQQCFHSEY FDSLLHACKP CHLRCSNPPA TCQPYCDPSV TSSVKGTYTV LWIFLGLTLV
LSLALFTISF LLRKMNPEAL KDEPQSPGQL DGSAQLDKAD TELTRIRAGD DRIFPRSLEY
TVEECTCEDC VKSKPKGDSD HFFPLPAMEE GATILVTTKT GDYGKSSVPT ALQSVMGMEK
PTHTR (SEQ ID NO: 2)

Fig. 7B
Mouse BCNA

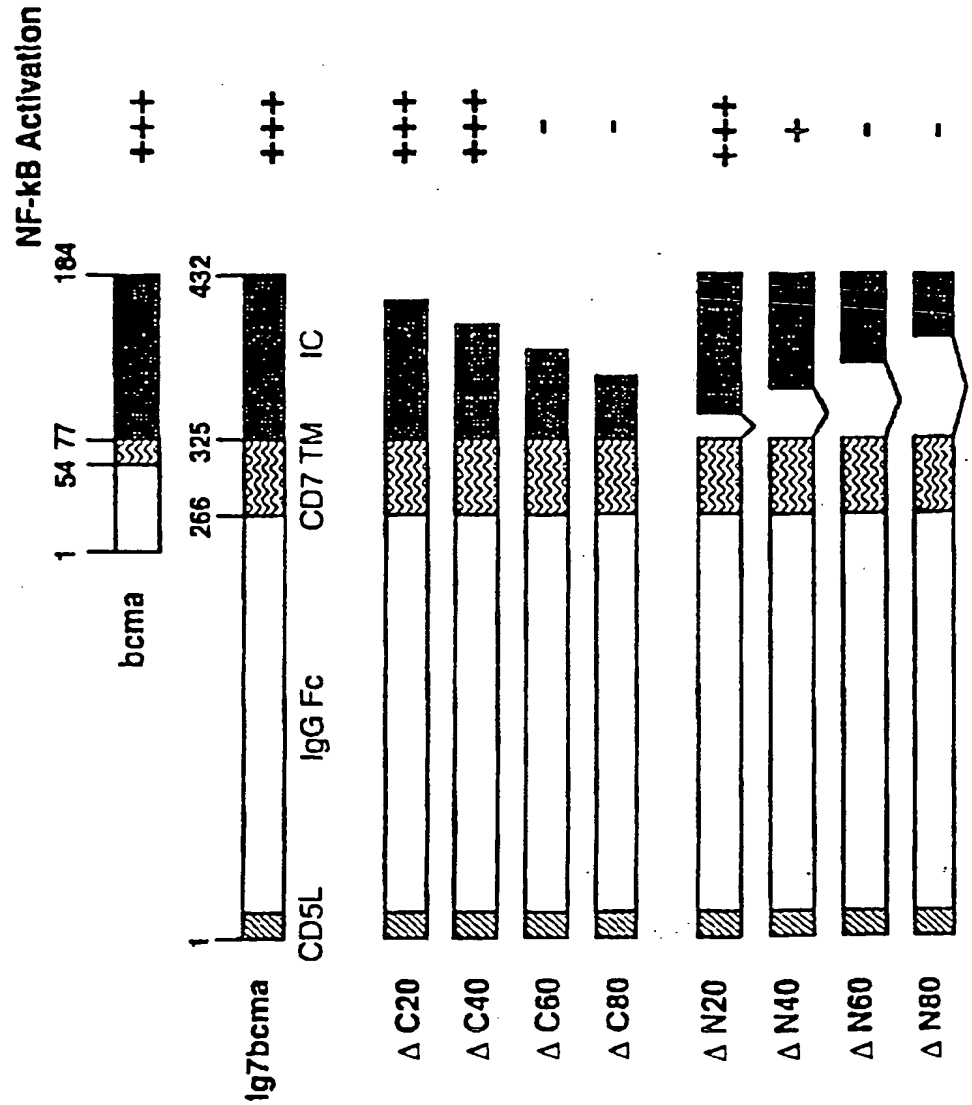


Fig. 8

09/914119

PATENT COOPERATION TREATY

PCT

REC'D 23 JUL 2001
WIPO PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 00786-371WO2	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US00/04925	International filing date *(day/month/year)* 24 FEBRUARY 2000	Priority date *(day/month/year)* 24 FEBRUARY 1999
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant THE GENERAL HOSPITAL CORPORATION		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 5 sheets.

   ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 2 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☒ Non-establishment of report with regard to novelty, inventive step or industrial applicability
   - IV ☐ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
05 SEPTEMBER 2000	13 JUNE 2001
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI TERRY J. DEY PATENT LEGAL SPECIALIST TECHNOLOGY CENTER 1600 Telephone No. (703) [illegible]

Form PCT/IPEA/409 (cover sheet) (July 1998) ★

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/04925

I. Basis of the report

1. With regard to the **elements** of the international application:*

[X] the international application as originally filed

[X] the description:
pages 1-29, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the claims:
pages 30-37, as originally filed
pages NONE, as amended (together with any statement) under Article 19
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the drawings:
pages 1-8, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the sequence listing part of the description:
pages NONE, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language ____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

[ ] contained in the international application in printed form.

[ ] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages NONE

[X] the claims, Nos. NONE

[X] the drawings, sheets/~~fig~~ NONE

5. [ ] This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

** Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

***Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US00/04925

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been and will not be examined in respect of:

☐ the entire international application.

☒ claims Nos. 5-6, 17-18, 24-43

because:

☐ the said international application, or the said claim Nos. _ relate to the following subject matter which does not require international preliminary examination *(specify)*.

☐ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. _ are so unclear that no meaningful opinion could be formed *(specify)*.

☐ the claims, or said claims Nos. _ are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for said claims Nos. 5-6, 17-18, 24-43 .

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1-4, 7-16, 19-23	YES
	Claims	NONE	NO
Inventive Step (IS)	Claims	NONE	YES
	Claims	1-4, 7-16, 19-23	NO
Industrial Applicability (IA)	Claims	1-4, 7-16, 19-23	YES
	Claims	NONE	NO

2. citations and explanations (Rule 70.7)

Claims 1-4, 7-16, 19-23 lack an inventive step under PCT Article 33(3) as being obvious over either Birnbaum (Journal of Virology, April 1994, Vol 68, No. 4, pages 2521-2528) or Clem et al (Science, vol. 254, November 1991, pages 13881390)in view of Zolotukin et al (US Patent 5,874,304).

Birnbaum et al teach an apoptosis -inhibiting gene from s nuclear polyhedrolsis virus encoding a polypeptide with Cys/His sequence motifs. The reference teaches that two different baculovirus genes are known to be able to block apoptosis triggered upon infection of Spodoptera cells with p35 mutants of the insect baculovirus. the reference teaches that using a genetic complementation assay to identify additional genes which inhibit apoptosis during infection with p35 mutant, Birnbaum et al have isolated a gene OpNPV that was able to functionally substitute for AcMNPV p35. The reference teaches a genetic complementation assay to identify apoptosis blockin genes in which SF-21 cells were cotransfected with vAcAnh and the test baculovirus DNA and transfected cells monitored.

Clem et al teach prevention of Apoptosis by a Baculovirus Gene during Infection of Insect cells. Clam et al have identif ied annihilator (vAcAnh) viral mutant was identified from expression vectors. The mutant caused premature death of Spodoptera frugiperda cells.

Either Birnbaum et al or Clem et al have taught the method of identification of the gene or the polypeptide involved in apoptosis using the a reporter gene expression. However, Zolotukhin et al teach humanized green fluorescent protein genes and methods of use of the protein in several assays. The reference teaches that the expression vectors may comprise a multiple cloning sites that is operatively positioned downstream from gfp gene sequence, and these vectors are useful in addition to the uses in creating C terminal fusion proteins by cloning a second protein encoding DNA segment into the multiple cloning site so that it is contiguous and in frame with the gfp sequence. The reference teaches that the recombinant host cells will express GFP to produce encoded GFP protein, preferably in amount sufficient to allow GFP detection by its fluorescence. The reference teaches that the expression vectors comprising a GFP (Continued on Supplemental Sheet.)

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

CLASSIFICATION:

The International Patent Classification (IPC) and/or the National classification are as listed below:

IPC(7): C12Q 1/00, 1/70, 1/66; C12N 7/00, 5/00, 5/02; G01N 33/554; A61K 38/00; C07H 21/02, 21/04, 21/00 and US Cl.: 435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):

gene operatively linked to a selected gene, and the fusion protein being produced in amount sufficient to allow cell detection by detecting the green fluorescence of GFP. Thus, it would have been obvious to use the GFP protein as a reporter gene in the method of identifying the gene involved in the apoptosis.

Claims 1-4, 7-16, 19-23 meet the criteria set out in PCT Article 33(2) and (4), because the prior art does not teach the method of identifying the a polypeptide which increases or decreases gene expression from a promoter using a GFP reporter gene operably linked to said promoter.

---------- NEW CITATIONS ----------

NONE

09191411 9 /

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 August 2000 (31.08.2000)

PCT

(10) International Publication Number
WO 00/50633 A1

(51) International Patent Classification[7]: C12Q 1/00, 1/70, 1/66, C12N 7/00, 5/00, 5/02, G01N 33/554, A61K 38/00, C07H 21/02, 21/04, 21/00

(71) Applicant *(for all designated States except US)*: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).

(21) International Application Number: PCT/US00/04925

(22) International Filing Date: 24 February 2000 (24.02.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/121,485 24 February 1999 (24.02.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/121,485 (CIP)
Filed on 24 February 1999 (24.02.1999)

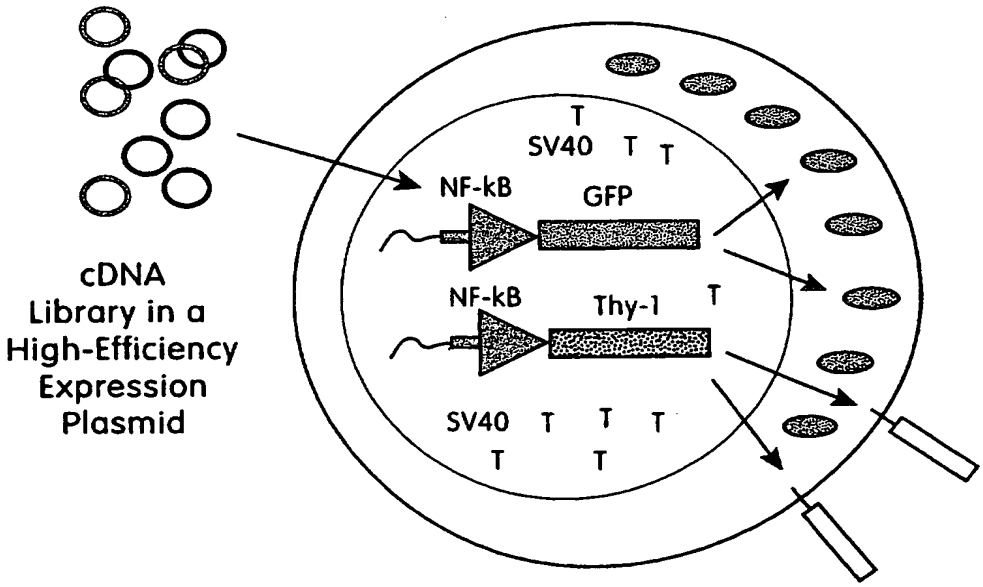
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SEED, Brian [US/US]; Apartment 5J, Nine Hawthorne Place, Boston, MA 02114 (US). TING, Adrian [MY/US]; 1675 York Avenue #6G, New York, NY 10128 (US).

(74) Agent: ELBING, Karen, L.; Clark & Elbing, LLP, 176 Federal Street, Boston, MA 02110 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: METHOD FOR CLONING SIGNAL TRANSDUCTION INTERMEDIATES



(57) Abstract: The invention features a method of identifying a polypeptide which increases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter (as shown in the figure), whereby expression of the reporter gene is increased if the library includes a polypeptide which increases gene expression from the promoter; (b) determining whether the reporter gene expression is increased in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is increased, identifying a polypeptide of the library which increases reporter gene expression.

WO 00/50633 A1

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(48) **Date of publication of this corrected version:**
11 October 2001

(15) **Information about Correction:**
see PCT Gazette No. 41/2001 of 11 October 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD FOR CLONING SIGNAL TRANSDUCTION INTERMEDIATES

Statement as to Federally Sponsored Research

This research has been sponsored in part by grant number AI27849 from the National Institutes of Health. The U.S. government has certain rights to this invention.

Background of the Invention

The invention relates to methods of identifying polypeptides and compounds which regulate gene expression.

Pharmaceuticals have historically been developed by testing libraries of up to several thousand compounds in laboratory animals, usually one compound at a time. The slow pace of the process, and its unsuitability for screening large numbers of diverse compounds, led to the development of approaches based on assays that can be completed quickly and *ex vivo*. With these approaches, the pharmaceutical drug discovery process has evolved into a catenation of several, sometimes partially concurrent, phases.

In the first phase, target identification and validation, a candidate target for a potential drug is identified by various means. These means may include hypotheses formed from the study of the pathophysiology of disease in humans or experimental animals, analysis of candidate signal transduction pathways *in vitro*, natural experiments such as genetic disorders of humans or other animals, results from targeted or random gene disruptions in model organisms, or disclosures by competitors. The projected consequences of a hypothetical drug interacting with its intended target to stimulate, block or

modulate the target activity is then tested. If the outcome of testing suggests that a drug affecting the target would have the desired physiological consequences in an intact organism, the target is said to be validated. Neither target identification nor validation are formally required, and, indeed, the historical process had neither of these phases.

Present models for pharmaceutical industry productivity emphasize the importance of efficiently identifying large numbers of validated targets and of developing rapid screens for their activity. The impetus for this model is the widespread belief that targets cannot be predicted effectively, so that a large number of targets must be evaluated to develop a small number of drugs. Hence, there remains an important need across the industry to accelerate the pace of target discovery and validation.

Once the target has been identified and validated, an assay must be created to allow large libraries of synthesized or natural compounds to be tested for their ability to interact with the target. When a compound is identified which acts on the target specifically, it is usually said to be a hit. The definition of a hit generally encompasses compounds that pass various secondary tests to assure that their activity is specific to the target of interest. From the various hits that are generated, one or more structures are chosen to represent the starting point for a program of systematic modification of the chemical structure. These structures are called lead compounds, and they are frequently selected from hits on the basis of their compatibility with directed synthesis programs, expected toxicity, or expected absorption, distribution, metabolism or excretion characteristics. Again, the process encompasses a wide range of practices and although they are usually distinguished, hits and lead compounds are sometimes referred to in ways that make them appear equivalent.

Summary of the Invention

In general, the invention provides a novel, rapid method for identifying members of selected signal transduction pathways which are targets for drug design.

In a first aspect, the invention features a method for identifying a polypeptide which increases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter, whereby expression of the reporter gene is increased if the library includes a polypeptide which increases gene expression from the promoter; (b) determining whether the reporter gene expression is increased in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is increased, identifying a polypeptide of the library which increases reporter gene expression.

In a second aspect, the invention features a method for identifying a polypeptide which decreases gene expression from a promoter, including (a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to the promoter, whereby expression of the reporter gene is decreased if the library includes a polypeptide which decreases gene expression from the promoter;

(b) determining whether reporter gene expression is decreased in the cell as a result of contact with the polypeptide library; and

(c) if reporter gene expression is decreased, identifying a polypeptide which decreases reporter gene expression.

In a third aspect, the invention features a method for identifying a polypeptide which modulates activation of a transcription factor activation domain, including (a) contacting a library of polypeptides with a cell that

expresses a recombinant anti-cell death gene and a chimeric transcription factor consisting of a yeast or bacterial DNA binding domain fused to a mammalian transcription factor activation domain, and that contains a reporter gene operably linked to a promoter consisting of a basal promoter and binding sites for the DNA binding domain, whereby expression of the reporter gene is altered if the library comprises a polypeptide which modulates activation of the transcription factor activation domain; (b) determining whether reporter gene expression is altered in the cell as a result of contact with the polypeptide library; and (c) if reporter gene expression is altered, identifying a polypeptide which modulates reporter gene expression.

In a fourth aspect, the invention features a method for identifying a compound which modulates gene expression from a promoter, including (a) contacting a library of compounds with a cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to the promoter, whereby expression of the reporter gene is altered if the library includes a compound which modulates gene expression from the promoter;

(b) determining whether reporter gene expression in the cell is altered as a result of contact with the compound library; and

(c) if reporter gene expression is altered, identifying a compound from the library which modulates reporter gene expression.

In a fifth aspect, the invention features a method for identifying a compound which decreases gene expression, including (a) contacting a library of compounds with a cell expressing (i) a recombinant anti-cell death gene; (ii) a second gene encoding a polypeptide; and (iii) a reporter gene that would have decreased expression if the function of the polypeptide was blocked; (b) determining whether expression of the reporter gene is decreased as a result of contact with the compound library; and (c) if expression of the reporter gene is

decreased, identifying a compound from the library which decreases reporter gene expression.

In preferred embodiments of the first, second, and third aspects, a library of DNA molecules encoding the library of polypeptides are expressed in a cell. The cell can be the same cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter (in which case the polypeptide is produced by the same cell that expresses a recombinant anti-cell death gene and that contains a reporter gene operably linked to a promoter), or the polypeptide can be produced by a cell other than the cell that expresses a recombinant anti-cell death gene and that includes a reporter gene operably linked to a promoter; preferably the DNA molecules are expressed from a high-efficiency expression system.

In other preferred embodiments of the first, second, and third aspect, the library of DNA molecules is introduced to the cell by transfection, and the mean number of DNA molecules introduced by transfection to the cell is at least 25. Preferably, the mean number of DNA molecules introduced by transfection into the cell is at least 100 or even 500.

In still other preferred embodiments of the first, second, and third aspect, the polypeptide is selected from the group consisting of an extracellular ligand, a cell surface receptor, and a signal transduction intermediate, and the DNA molecules are expressed from a high-efficiency expression system.

In preferred embodiments of the first, second, third, fourth, and fifth aspects, step (c) includes (i) dividing the library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which activates reporter gene expression is identified; the promoter can derived from a mammal and the library of polypeptides can include polypeptides derived from a bacterium or a virus.

In other preferred embodiments of the first, second, third, fourth, and fifth aspects, the promoter is a heterologous promoter, the reporter gene is GFP, the anti-cell death gene is selected from the group consisting of bcl family members, IAP family members, and crmA, and the cell is selected from the group consisting of CHO, CD-1, Cos, 293, HeLa, BHK, or L cells.

In another aspect, the invention features a method for determining whether a compound modulates NF-kB biological activity, the method including the steps of: (a) providing a cell expressing a BCMA polypeptide; (b) contacting the cell with a candidate compound; and (c) measuring the level of expression of the BCMA polypeptide in the cell, wherein a change in the level of expression of the BCMA polypeptide in the cell, relative to a cell not contacted with the candidate compound, identifies the candidate compound as a compound that modulates NF-kB activity.

In still another aspect, the invention features a method for determining whether a compound modulates BCMA biological activity, the method including the steps of: (a) providing a BCMA polypeptide; (b) contacting the polypeptide with a candidate compound; and c) measuring the level of biological activity of the BCMA polypeptide, wherein a change in the level of biological activity of the BCMA polypeptide, relative to a polypeptide not contacted with the candidate compound, identifies the candidate compound as a compound that modulates BCMA activity. The BCMA polypeptide can be in a cell or in a cell-free system. A preferred BCMA biological activity is the modulation of NF-kB biological activity (e.g., the modulation of transcription by NF-kB) or NF-kB expression. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

In yet another aspect, the invention features a method for determining whether a compound modulates NF-kB activity, the method including the steps of: (a) providing a BCMA polypeptide; (b) contacting the polypeptide with a candidate compound; and (c) detecting the binding of the candidate compound to the polypeptide, wherein a candidate compound that binds to the polypeptide is a compound that modulates NF-kB biological activity. The BCMA polypeptide can be in a cell or in a cell-free system. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

In another aspect, the invention features a substantially purified polypeptide consisting of a BCMA polypeptide molecule lacking a BCMA extracellular domain. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2). Other preferred polypeptides are BCMA polypeptides that modulate NF-kB activity.

In yet another aspect, the invention features an NF-kB modulator including a BCMA polypeptide covalently linked to a heterologous compound. Preferably, the BCMA polypeptide includes a polypeptide sequence having substantial identity to amino acids 98 to 164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2). Preferably, the modulator modulates NF-kB biological activity or expression.

The invention also features methods for activating NF-kB activity in a cell by contacting the cell with a recombinant BCMA polypeptide having NF-kB activating activity or by contacting the cell with a recombinant nucleic acid molecule encoding a BCMA polypeptide having NF-kB activating activity.

In related aspects, the invention features the use of BCMA polypeptides and nucleic acids for preparing pharmaceutical compositions for treating cancer or apoptosis.

By a "reporter gene" is meant a DNA or RNA sequence which encodes a reporter protein that is capable of being readily detected either inside or outside a cell.

The reporter gene is operably linked to a promoter which shows low spontaneous activation, so that the activity of the reporter protein in the presence of the activating polypeptide be at least two standard deviations greater than the activity in its absence. Various methods of increasing the sensitivity of reporter genes are known in the art, including: deletion from natural genes of inhibitory sequences, which may be found both within and external to the transcribed portion and which may affect mRNA formation, stability, or translational efficacy; addition of efficiently utilized introns to increase the rate of formation of mature mRNA; multimerization of upstream activation regions or binding sites for known transcriptional activators; increasing the copy number of the reporter gene; and protection of the activity of the reporter gene from adventitious stimulatory or inhibitory activity in flanking DNA by inserting the reporter gene between matrix attachment regions or chromatin insulator sequences.

Many different types of reporter proteins are known in the art. They frequently comprise proteins not normally found, or present in minor amounts, in some cells; they include enzymes that detoxify antimicrobial agents, such as aminoglycoside or aminocyclitol phosphotransferases or acetyltransferases, beta-lactamases or chloramphenicol acetyltransferase; enzymes of diverse origin that catalyze chromogenic, fluorogenic, or chemiluminescent reactions in the presence of exogenous substrates, such as beta-galactosidase, beta-

glucuronidase, alkaline phosphatase, catechol 2,3-dioxygenase, or various peroxidases; enzymes that catalyze photoreactions, such as bacterial or firefly luciferases; enzymes, like glycosyl transferases, that generate nonproteinaceous structures easily detected by antibodies, lectins, or cognate binding proteins; proteins easily monitored upon cell surface expression or secretion such as surface or secreted antigens for which corresponding antibodies or recognition proteins are known; and proteins which catalyze the synthesis of, or stoichiometrically embody, fluorescent structures, without exogenous substrates, such as the jellyfish fluorescent proteins (e.g., GFP).

By "operably linked" is meant that the gene for the reporter protein is positioned adjacent to a promoter which directs transcription of the gene and, ultimately, facilitates expression of the reporter protein.

By "promoter" is meant any minimal nucleic acid sequence sufficient to direct transcription of the reporter gene. The promoter is one which is activated by binding to a polypeptide. Examples of promoters useful in the invention are promoters which are normally linked to genes which are expressed when a cell is in a pathologic or disease state (e.g., cancer, inflammation, or due to bacterial or viral infection), and the protein products of which are directly or indirectly responsible for this state. Suitable promoters include, but are not limited to, the NF-kB promoter, the interleukin-2 promoter, and the HIV-1 long terminal repeat promoter.

A library of DNA molecules refers to a set of DNA molecules, each in a DNA expression vector. Preferably, the DNA expression vector displays high efficiency such that the level of expression is high. While in most cases, the library includes DNA molecules encoding tens, hundreds, or even thousands of different polypeptides, DNA molecules in a library can also encode only one polypeptide (for example, during the final steps of a sib

selection). A library with hundreds of different DNA molecules is considered to have greater "complexity" than a library with five different DNA molecules. Like DNA libraries, compound libraries can have different degrees of complexity. One feature of the invention is a method which allows for following an activity of interest through the sequential screening of libraries with less and less complexity.

By "candidate compound" is meant a chemical, be it naturally-occurring or artificial, that is surveyed for its ability to modulate BCMA or NF-kB biological activity. Candidate compounds may include, for example, peptides, polypeptides, antibodies (and fragments thereof), synthesized organic molecules, naturally occurring organic molecules, nucleic acid molecules, and components or derivatives thereof. The candidate compounds may be screened using any of the methods described herein using an animal, a cell, a lysate or extract derived from a cell, or a molecule derived from a cell. The measuring may be, for example, for the purpose of detecting altered gene expression, altered RNA stability, altered protein stability, altered protein levels, altered protein phosphorylation, or altered protein biological activity. The means for measuring may include, for example, antibody labeling, immunoprecipitation, phosphorylation assays, and methods known to those skilled in the art for detecting nucleic acids.

By "modulating" is meant changing, either by decrease or increase.

By "BCMA polypeptide" is meant a polypeptide that has substantial identity to human or mouse BCMA as shown in Figs. 7A and 7B, respectively, over a region of twenty consecutive amino acids and has a BCMA biological activity.

By "substantial identity" is meant that two polypeptide sequences, when optimally aligned, such as by the GAP or BESTFIT programs using

default gap weights, share at least 80% sequence identity, more preferably at least 90% sequence identity, and most preferably at least 95% or even 99% sequence identity or more.

By "heterologous compound" is meant a polypeptide, chemical, or detectable label (e.g., biotin) that is not naturally associated, in this case, with the polypeptide.

By "BCMA biological activity" is meant BCMA-mediated modulation of NF-kB biological activity, or expression or the binding of an antibody that specifically binds a BCMA polypeptide.

By "NF-kB biological activity" is meant any function performed by activated NF-kB.

The invention features a method of screening DNA and compound libraries for their ability to modulate reporter gene expression in a cell which is expressing a recombinant anti-cell death gene. The expression of this gene allows the cell to survive in conditions which would otherwise lead to its death and, as a result, failure of the screening procedure.

This method is suitable for identifying polypeptides that modulate transcription from a selected promoter. These polypeptides are, in turn, targets for drugs. The method is also suitable for identifying compounds which either mimic or block function of a polypeptide which itself modulates transcription from a selected promoter.

Hence, the invention features a new, efficient multistep method for identifying lead compounds which modulate expression from a promoter for development of pharmaceutical compounds.

The invention also features a new NF-kB activator, BCMA. As NF-kB is involved in numerous cellular processes and disease states, BCMA is useful for the treatment of disease. BCMA is also useful for the identification

of compounds that modulate its expression or biological activity. These compounds may be developed as drugs, or used as lead compounds for the purpose of identifying drugs.

Other features and advantages of the invention will be apparent from the following detailed description, the drawings, and the claims.

Brief Description of the Drawings

Fig. 1 is a schematic illustration of a reporter cell system for rapid identification of cDNAs encoding polypeptides of interest. A cDNA library is transfected into a cell which includes a reporter gene (e.g., GFP or Thy-1) operably linked to a promoter (in this case the NF-kB promoter). If the polypeptide expressed by one of the plasmids in the reporter cell is capable of inducing expression from the NF-kB promoter, the reporter gene is expressed, and the detectable marker protein is produced.

Fig. 2 is a schematic illustration of a cDNA library being divided into smaller libraries. Each of the libraries is then transfected into reporter cells. Libraries which induce reporter gene expression are then further divided until the cDNA(s) encoding the polypeptide which induces reporter gene expression are isolated.

Fig. 3 is a schematic illustration of three stages of enrichment for an intracellular signal transduction intermediate. The ratios represent the occurrence of the cDNA encoding the reporter gene-inducing polypeptide compared to the total cDNAs. Note that, during transfection, each cell receives more than one plasmid.

Fig. 4 is a schematic illustration showing how secreted factor can be distinguished from intracellular signal transduction intermediates such as receptors or kinases. Soluble ligands diffuse through the medium to engage

receptors on untransfected cells, which can lead to reporter gene expression in most, if not all, cells. Intracellular effectors are restricted to the transfected cells.

Fig. 5 is a schematic illustration of arrayed clones. N^2 elements, pooled into N columns and N rows, can be assayed with 2N transfections. In the example provided, 4 rows and 4 columns were positive. 4^2 or 16 transfections are then required to identify the intersections unambiguously.

Fig. 6 is a schematic illustration of an indexed library protocol. The method is similar to the one described in Fig. 1 to Fig. 5, except that individual plasmids are grown as individual cultures. Libraries of plasmids are then prepared and assayed. This method facilitates the rapid recovery of individual plasmids, and prevents the loss of positives when a library is divided into libraries of lower complexity.

Fig. 7A is a schematic illustration showing human BCNA polypeptide sequence (GenBank accession number Q02223). The putative transmembrane domain is indicated in bold.

Fig. 7B is a schematic illustration showing mouse BCNA polypeptide sequence (GenBank accession number AAC23799).

Fig. 8 is a schematic illustration showing that the intracellular domain of human BCMA was capable of activating NF-kB. A series of fusion proteins were constructed using the CD5L leader sequence, IgG Fc, a CD7 transmembrane domain, and amino acids from human BCMA (SEQ ID NO: 1) as follows: Ig7bcma--amino acids 78-184; ΔC20--amino acids 78-164; ΔC40--amino acids 78-144; ΔC60--amino acids 78-124; ΔC80--amino acids 78-104; ΔN20--amino acids 98-184; ΔN40--amino acids 118-184; ΔN60--amino acids 138-184; ΔN80--amino acids 158-184.

Detailed Description

The present invention provides a method for the rapid identification of molecules involved in regulation of specific signal transduction pathways. It can be used both as a target identification tool and as a rapid assay for drug effectiveness, and as such can significantly reduce the amount of time needed to go from target to hit.

The method features an expression cloning approach that identifies polypeptides that have the ability to activate reporter genes (Fig. 1). Such polypeptides are called here activating polypeptides. Expression cloning is a technique which identifies polypeptides solely on the basis of their ability to generate an observable activity of interest. In order for that activity to be identified, the polypeptide must generally be expressed, either *in vivo* or *in vitro*, and a suitably sensitive assay must be available to detect the activity after the polypeptide has been expressed. In the present method, it is preferred that the activity that is detected is dependent on the output of a reporter gene (Fig. 1).

Although the use of expression cloning in conjunction with reporter genes is not new, existing methods of application have not resulted in highly efficient systems for the identification of large numbers of new molecules. The present approach is a high throughput system for the identification of cDNA clones encoding polypeptides that induce reporter gene activity. Some features of the high throughput system of the present invention are: (1) a very high efficiency cDNA expression plasmid; (2) an easily detected reporter molecule; (3) features to prevent the death of the reporter cell due to toxicity of expressed genes; and (4) a method for the introduction of large numbers of plasmids into multiple cells.

Expression cloning can be carried out by either of two general

paradigms, positive selection, or sib selection, also known as pool division. By positive selection is meant a method for the enrichment of cells, viruses or genetically linked assemblies of proteins and the nucleic acids that encode them by physically separating the cells or genetically linked assemblies from the much larger collection of cells or assemblies that encode molecules of little interest. Because the nucleic acids that encode the polypeptides of interest are physically linked to the cells, viruses, or other assemblies that they comprise, the nucleic acid is physically purified by positive selection and the process can be repeated until a single nucleic acid is found that encodes the activity of interest. It is desirable in positive selection schemes to develop methods for association of a single nucleic acid with a single protein assembly or cell. Thus in positive selection it is desirable to identify and use methods for the introduction of nucleic acids into cells that result in a single nucleic acid species per cell.

The method of the invention uses sib selection, a process of screening in which pools of molecules that are derived from cognate libraries of nucleic acids are assayed for activity, and positive libraries are detected by some signal (e.g., expression of a reporter gene). The nucleic acids that make up the library are then separated into libraries with less complexity, which are then reassayed and redivided until a single nucleic acid is found that encodes the polypeptide with the desired activity (Figs. 2, 3, and 4) Because sib selection depends on the detection of the activity of aggregates of nucleic acids, it is often advantageous to use methods for the transfection or introduction of nucleic acids into cells that result in a large number of nucleic acid species per cell. Sib selection schemes, like positive selection methods, can also be carried out entirely *in vitro*.

Features of the Invention

Transfection Method

For sib selection to be an efficient method of cloning genes, according to the invention, it is important that a large number of DNA molecule are introduced into each cell. Many methods of introducing DNA molecules are known in the art, including microinjection, complexation with positively charged synthetic polycations such as DEAE dextran, polybrene, polylysine, or polyarginine, complexation with histones and other basic proteins, complexation with cationic lipids or related amphipathic molecules, condensation with polyethylene glycol or polyhydroxybutyrate, coprecipitation with calcium phosphate, electroporation, scrape loading or partial rupture, and fusion with bacterial or microbial spheroplasts. Preferred among these are methods that can be easily carried out in parallel and that result in the co-introduction into the reporter cell of multiple nucleic acids that encode different species. Moreover, the method of transfection will preferably provide, on average, at least 25 DNA molecules per cell. More preferably, the mean number will be at least 100 or even 500 DNA molecules per cell. The methods which achieve the preferred results include calcium phosphate coprecipitation, complexation with polycations or cationic lipids, and condensation with uncharged polymers such as polyethylene glycol.

The libraries of nucleic acid molecules encoding potentially activating polypeptides can also be created in a biologically active assembly, such as a virus or viral transducing particle, which is capable of introducing itself into the reporter cell directly. In such a case a cDNA library is prepared in the viral vector, and libraries of active virus or transducing particles are applied to the reporter cell. In general the reporter cell will have been previously engineered to contain a reporter gene, but the reporter gene may also

be introduced concurrently with the activating nucleic acid.

High Efficiency cDNA Expression Vector

To detect ectopic activation of a signal transduction pathway it is desirable to greatly overexpress proteins that act as signal transduction intermediates.

The invention uses a high efficiency cDNA expression system to produce proteins in the reporter cell. Such a system is provided by the use of a very strong promoter, such as, for example, the elongation factor 1 alpha (EF-1α) promoter; 3' untranslated region (3' UTR) and polyadenylation consensus (poly(A)) sequences from the human growth hormone gene; and the human IgG1 hinge-CH2 intron. Other strong promoter and nonpromoter elements are known in the art (for example, the murine or human cytomegalovirus immediate early gene promoters, globin introns, and 3' UTR/poly(A) sequences).

Cell Death Inhibition

Another important feature of the present system is a method to prevent the death of cells overexpressing proteins. Such a method is important for two reasons: signal transduction intermediates themselves can lead to cell death if they are expressed at high levels, and, in addition, if libraries of nucleic acids are transfected, the presence of even a low frequency of nucleic acids encoding toxic proteins can interfere with detection of the desired signal. The latter effect can be predicted to have greater impact as the size of the library increases.

There are several known methods to prevent the demise of cells undergoing programmed cell death, or apoptosis. Both viral and cellular

antagonists of cell death are known, and among them are species that act upstream in the apoptosis pathway, or at multiple levels. Examples of the former include dominant negative forms of cell death proteins, such as FADD or TRADD, and cellular inhibitors of apoptosis, such as members of the viral or cellular IAP family. Examples of proteins that act at multiple levels, or relatively downstream, include the orthopoxvirus CrmA and baculovirus p35 proteins, members of the cellular Bcl family, and both peptide and nonpeptide inhibitors of caspases, the cysteine proteases which act in a zymogen cascade to generate the death program. For the prevention of apoptosis in expression cloning, combinations of broad spectrum antiapoptotic proteins with different mechanisms of action are preferred, such as CrmA and Bcl-xL.

Examples

Identification of gene products leading to the activation of the transcription factor NF-kB

We established a reporter cell line by stably transfecting the commercially-available human embryonic kidney cell line, 293 EBNA, with two constructs: a reporter construct consisting of NF-kB promoter elements upstream of the green fluorescent protein (GFP); and a eukaryotic expression construct that expresses two anti-cell death molecules, CrmA and Bcl-xL. The former functions as a readout for the presence of signals that activate the NF-kB signal transduction pathway, whereas the latter prevents these cells from undergoing programmed cell death in the event that pro-apoptotic signals are present. The expression in the reporter cell line of NF-kB-activating molecules, including receptors (e.g. tumor necrosis factor-α receptor 1), ligands (e.g. tumor necrosis factor-α) and intracellular signaling proteins (e.g. RIP), results in the robust production of GFP. The level of GFP production, as a

measure of NF-kB activation, can be assayed using a fluorescent microscope or flow cytometry. High level expression of these various NF-kB-activating molecules was achieved by placing their cDNAs under the control of the very strong EF-1α promoter present in the commercially available plasmid PEAK8 (Edge Biosystems). Thus, one can transfect into these cells cDNA libraries and identify individual cDNAs which are capable of NF-kB activation.

A cDNA library, prepared from activated human T cell mRNA and cloned into the PEAK8 vector, was subdivided into smaller libraries of approximately 500 cDNA clones each, and DNA was prepared from each library. The DNA from the libraries were transfected into the reporter cell line by calcium phosphate precipitation following published protocols, e.g., Ausubel et al. et al., 1997, *Current Protocols in Molecular Biology*, Wiley Interscience. Cells were allowed to continue to grow for 48 hours and assayed for GFP production. cDNA libraries that gave a positive GFP signal, as defined by the appearance of at least 0.1% of bright cells by fluorescence microscopy, were screened further by sub-dividing into libraries of 50 cDNAs each (Fig. 2). DNA from each library was prepared, transfected into the reporter cell line and assayed for GFP production. The process was repeated until a single cDNA clone that induced a positive GFP readout was obtained. The screening procedure has resulted in the cloning of DNAs encoding known NF-kB-activating molecules such as surface receptors (DR3, FAS), soluble ligands (interleukin-1, TRAIL, CD40 ligand), intracellular signaling molecules (small molecular weight GTPase rho). Also identified were a previously known molecule having no known function (BCMA) and novel molecules.

The inclusion of two anti-cell death genes in the transfected cells are likely to have aided in the expression cloning of at least some of the above-mentioned DNAs. It has been established that expression of either FAS or DR3

in a cell would induce apoptosis in that cell. The fact that each was retrieved substantiates the improved nature of the expression cloning method of the present invention. Similarly, as the method of the present invention preferably employs high copy number, it is likely that expression of anti-cell death genes allows for cells that otherwise would have died, due to a high number of inserts, to survive.

The method is more rapidly and efficiently applied by using indexed arrays of bacterial cultures, in which each culture is derived from a single colony, and, thus, represents an independently derived cDNA expression plasmid (Fig. 5). By pooling small amounts of the cultures from rows and columns of large arrays and preparing DNA from each of the libraries, it is possible to assess the relative activity of every row and every column, thereby reducing labor dramatically (Fig. 6). For example if there are 90,000 individual/clones arrayed in a 300 by 300 matrix, the 300 rows and 300 columns can be transfected to determine all of the rows and all of the columns that bear positive clones. Hence the array can be indexed in only 600 DNA preparations and transfections. If there is only one positive culture, the job is complete. If there are multiple positive cultures, then in the worst case, each row and each column will have only one positive culture, and if there are x positives, then there are x^2 possible intersections. As long as x is a small number, though, the work involved is quite modest. For example if x were 10, the entire array of 90,000 clones could be screened in only 700 DNA preparations and transfections.

Variations of the experimental approach outlined above are applicable depending on the particular system that is to be examined. The reporter system and cell lines can be adapted to the promoter to be investigated. For instance, the NF-kB promoter in the GFP reporter construct can be replaced

by the interleukin-2 promoter, introduced into a T lymphocyte cell line and screened for molecules that regulate the signal transduction pathway leading to the expression of interleukin-2 in T lymphocytes. Other promoters can be drawn from viral sources, such as the HIV-1 long terminal repeat (LTR) promoter, or other inducible promoters of viral origin. The calcium phosphate transfection procedure can be changed to another method more suitable for the particular reporter cell type. The alternative transfection protocols could be based on electroporation, cationic lipids, DEAE-dextran, spheroblasts fusion or viral-mediated delivery.

A variation to the method described above is to screen for molecules that will turn off expression from a promoter. For instance, the NF-kB reporter in the reporter cell line is activated in response to interleukin-1 stimulation. By including interleukin-1 in the culture medium, one can then search for molecules that will inhibit the interleukin-1 dependent expression from the NF-kB promoter. Depending on the cDNA library that is used, these molecules can be either ones that are naturally negative regulatory or they can be mutant versions that behave in a dominant inhibitory manner. Similarly, screens can be conducted with a combinatorial library to look for small pharmacological molecules that will negatively interfere with the pathway (see below).

The screening methods described above are well-suited for screening for genes from one organism that interact with a pathway in another organism. A prime example is screening the expressed genomes of viruses, bacteria or other pathogens for genes that, when expressed, might interact with the NF-kB pathway, This method, utilized by these pathogens to alter the immune response to their advantage, can identify potential targets for pharmacological interventions.

The reporter cell may also be provided with proteins which increase

the efficiency of the reporter gene. These proteins are usually introduced by transfection of an expression plasmid that encodes them. Proteins which increase the efficiency of the reporter gene may provide essential elements of a signaling pathway which are not otherwise present in the reporter cell, or may consist of artificial proteins that act to integrate, amplify, or selectively respond to signals from the pathway of interest. Of special relevance are artificial transcriptional activators which consist of a DNA binding element that interacts specifically with cognate binding sites in an artificial promoter and a pathway-specific transcriptional activator element that responds to activation of the pathway with a change in transcriptional activity. Such artificial activators are known in the art as fusions between bacterial or yeast DNA binding proteins and mammalian transcription factor activation domains. Such activation domains provides sites for protein binding, ligand-activated conformational change or post-translational modification, that increase the transcription-promoting capabilities of the artificial activator. Examples of artificial activators that are commercially available or known in the art include LexA and Gal4 fusions with c-Jun, Elk1, CREB, c-Fos, ATF2, CHOP, and members of the nuclear hormone superfamily.

In one example, a cell which contains (i) a recombinant anti-cell death gene, (ii) a chimeric transcription factor consisting of the Gal4 DNA binding domain fused to the c-Jun activation domain, and (iii) DNA encoding GFP operably linked to regulatory sequence consisting of a basal promoter and Gal4 binding sites is used to identify polypeptides or compounds which modulate c-Jun activation. If, for example, expression of a polypeptide leads to c-Jun phosphorylation (i.e., activation), then increased GFP expression would result.

BCMA is an activator of NF-kB

Using the methods described herein, we identified BCMA as an NF-kB activator. BCMA was discovered by molecular analysis of a t(4;16) translocation characteristic of a human T cell lymphoma (Laabi et al., EMBO J. 11: 3897-3904, 1992), and its function was not previously known.

Based on the present discovery, BCMA is a target for drug discovery or rational drug design. For example, a compound that modulates BCMA expression or biological activity will also modulate NF-kB biological activity. Accordingly, the invention features methods and reagents for the identification of NF-kB modulating compounds.

BCMA polypeptides or nucleic acid molecules are also useful for the treatment of diseases associated with insufficient or inappropriate NF-kB biological activity or expression. BCMA polypeptides or nucleic acid molecules are administered to a patient using an appropriate delivery vehicle, as known in the art. Generally, the BCMA polypeptide or nucleic acid molecule is delivered in a pharmaceutically acceptable carrier.

Coupled Target Identification and Assay Generation

Directly relevant to the NF-kB signal transduction pathway is the identification of molecules that could potentially play a role in regulating inflammation and oncogenesis. The search for NF-kB activating molecules do not have to be restricted to screening cDNA expression libraries; the same principles can be used to identify compounds which modulate the output of the pathway, either by mimicking the activity of a polypeptide or by blocking its activity when the polypeptide is overexpressed. The latter compounds will act either upon or downstream of the overexpressed protein; once enough activating proteins have been identified, it will be possible to identify at what

step the compounds act by evaluating their action on a panel of transfected cells expressing different activating proteins.

In general, compounds are identified from large libraries of both natural product and synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (e.g., semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceangraphics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, e.g., by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (e.g., taxonomic

dereplication, biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known to modulate the test promoter should be employed whenever possible.

When a crude extract is found to modulate reporter gene expression, further fractionation of the positive lead extract is necessary to isolate chemical constituents responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having the desired activity. Methods of fractionation and purification of such heterogenous extracts are known in the art. If desired, compounds shown to be useful agents for the modulation of gene expression or biological activity are chemically modified according to methods known in the art.

Primary screens for compounds that modulate BCMA biological activity

Modulating NF-kB expression or biological activity modulates numerous important cellular processes. The finding that NF-kB activity is regulated by BCMA allows us to provide assays for drugs that modulate NF-kB by monitoring BCMA expression or biological activity. Such assays may measure BCMA expression by measuring changes in: (a) levels of BCMA protein; (b) levels of BCMA RNA; (c) levels of BCMA-mediated NF-kB biological activity; or (d) levels of a reporter gene or protein expressed from a NF-kB promoter. These measurements may be made *in vitro* or *in vivo*. These assays allow for the identification of compounds that modulate NF-kB biological activity (e.g., gene transcription). Such identified compounds may have therapeutic value, for example, in the treatment of diseases that result in too little or too much cell death.

Cells overexpressing BCMA can be produced using standard techniques. Compounds that are identified may bind to BCMA and prevent BCMA activation of NF-kB. While screening of compounds can be performed with cultures of primary cells, cell lines be also be used. Cell lines can be modified such that the cells constitutively express a BCMA polypeptide, for example, the BCMA intracellular domain.

Any cell line, such as ones described above, can also be engineered to contain a reporter gene expressed under control of the NF-kB promoter (described above). A preferred reporter gene codes for GFP. Typically, the expression of the gene (e.g., the endogenous NF-kB gene or a recombinant reporter gene expressed under the control of the NF-kB promoter or fragment thereof) is measured by assaying the RNA or protein levels or both of the expressed gene. For example, the polypeptide expressed by the NF-kB gene or by the reporter gene produces a detectable signal under conditions that allow compound-mediated changes to be measured. Quantitatively determining the amount of signal requires comparing the amount of signal produced in the absence of any compound being tested to the amount produced when the cell is contacted with the compound, as is described herein. The comparison permits the identification of the compound as one that causes a change in the detectable signal produced by the expressed gene (e.g., at the RNA or protein level) and thus identifies a compound that is capable of modulating NF-kB expression. In order to prevent the NF-kB cells from dying, a second gene encoding an apoptosis inhibitor can also be expressed in the cells, as described herein.

Secondary screens for compounds that modulate NF-kB activity

After test compounds that appear to modulate NF-kB expression are identified, it may be necessary or desirable to subject these compounds to

further testing. The invention provides such secondary confirmatory assays. For example, a compound that appears to modulate NF-kB activity in early testing may be subject to additional assays to confirm that the compound also modulates NF-kB activity *in vivo*. In the first round of *in vivo* testing, NF-kB activity is initiated in animals by well-known methods and the compound is administered by one of the means described in the "Therapy" section, immediately below. Cells or cellular tissue are isolated within hours to days following the insult, and are subjected to assays to assess the level of NF-kB expression or biological activity. Such assays are well known to those skilled in the art. Examples of such assays include, but are not limited to, ELISAs, Western blot analysis, RT-PCR, RIA, and Northern blot analysis.

Therapy

NF-kB is an important regulator of inflammatory responses (e.g., rheumatoid arthritis, inflammatory bowel disease, septic shock), apoptosis, oncogenesis, and anti-viral and anti-bacterial responses. Therefore, the discovery of new gene products that regulate NF-kB activity, and thus the disease process, will result in the identification of molecular targets for pharmacological intervention. By increasing or mimicking BCMA biological activity, one could, for example, boost anti-tumor antibody production or increase T cell cytotoxicity against tumor cells in cancer immunotherapy. Conversely, antagonizing BCMA biological activity would be advantageous, for example, in situations in which it is desirable to down-regulate immune cell function. Compounds, identified using any of the methods disclosed herein, may be administered to patients or experimental animals with a pharmaceutically- acceptable diluent, carrier, or excipient, in unit dosage form. Conventional pharmaceutical practice may be employed to provide suitable

formulations or compositions to administer such compositions to patients or experimental animals. Although intravenous administration is preferred, any appropriate route of administration may be employed, for example, parenteral, subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular, intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral administration. Therapeutic formulations may be in the form of liquid solutions or suspensions; for oral administration, formulations may be in the form of tablets or capsules; and for intranasal formulations, in the form of powders, nasal drops, or aerosols.

Methods well known in the art for making formulations are found in, for example, Remington: The Science and Practice of Pharmacy, (19th ed.) ed. Gennaro AR., 1995, Mack Publishing Company, Easton, PA. Formulations for parenteral administration may, for example, contain excipients, sterile water, or saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, or hydrogenated napthalenes. Biocompatible, biodegradable lactide polymer, lactide/glycolide copolymer, or polyoxyethyl ene-polyoxypropyl ene copolymers may be used to control the release of the compounds. Other potentially useful parenteral delivery systems for antagonists or agonists of the invention include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation may contain excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or may be oily solutions for administration in the form of nasal drops, or as a gel.

As described herein, we have discovered that BCMA activates NF-kB activity. NF-kB, in turn, activates numerous cellular processes. Hence, any compound that modulates NF-kB expression is a candidate compound for use

in drug development. One possible compound is a polypeptide fragment of BCMA which maintains its ability to bind to another component of the signal transduction pathway but has lost its ability to activate NF-kB. Such a polypeptide will act as an inhibitor of wild-type NF-kB signaling . Another possible compound is a polypeptide fragment of BCMA which exhibits constitutive activation of NF-kB.

Other Embodiments

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication or patent application was specifically and individually indicated to be incorporated by reference.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure come within known or customary practice within the art to which the invention pertains and may be applied to the essential features hereinbefore set forth.

What is claimed is:

1. A method for identifying a polypeptide which increases gene expression from a promoter, said method comprising the steps:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is increased if said library comprises a polypeptide which increases gene expression from said promoter;

(b) determining whether said reporter gene expression is increased in said cell as a result of contact with said polypeptide library; and

(c) if said reporter gene expression is increased, identifying a polypeptide of said library which increases said reporter gene expression.

2. The method of claim 1, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which activates reporter gene expression is identified.

3. A method for identifying a polypeptide which decreases gene expression from a promoter, said method comprising:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is decreased if said library comprises a polypeptide which decreases gene expression from said promoter;

(b) determining whether said reporter gene expression is decreased in said cell as a result of contact with said polypeptide library; and

(c) if reporter gene expression is decreased, identifying a polypeptide

which decreases said reporter gene expression.

4. The method of claim 3, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which decreases reporter gene expression is identified.

5. A method for identifying a polypeptide which modulates activation of a transcription factor activation domain, said method comprising:

(a) contacting a library of polypeptides with a cell that expresses a recombinant anti-cell death gene and a chimeric transcription factor consisting of a yeast or bacterial DNA binding domain fused to a mammalian transcription factor activation domain, and that comprises a reporter gene operably linked to a promoter consisting of a basal promoter and binding sites for said DNA binding domain, whereby expression of said reporter gene is altered if said library comprises a polypeptide which modulates activation of said transcription factor activation domain;

(b) determining whether said reporter gene expression is altered in said cell as a result of contact with said polypeptide library; and

(c) if reporter gene expression is altered, identifying a polypeptide which modulates said reporter gene expression.

6. The method of claim 5, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a polypeptide which modulates reporter gene expression is identified.

7. The method of claim 1, 3, or 5, wherein said contacting comprises expressing a library of DNA molecules in a cell, wherein said library of DNA molecules encodes said library of polypeptides.

8. The method of claim 7, wherein said cell is the same cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

9. The method of claim 7, wherein said library of DNA molecules is introduced to said cell by transfection.

10. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 25.

11. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 100.

12. The method of claim 9, wherein the mean number of said DNA molecules introduced by transfection into said cell is at least 500.

13. The method of claim 7, wherein said DNA molecules are expressed from a high-efficiency expression system.

14. The method of claim 1, 3, or 5, wherein said polypeptide is produced by the same cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

15. The method of claim 1, 3, or 5, wherein said polypeptide is produced by a cell other than the cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter.

16. The method of claim 1, 3, or 5, wherein said polypeptide is selected from the group consisting of an extracellular ligand, a cell surface receptor, and a signal transduction intermediate.

17. A method for identifying a compound which modulates gene expression from a promoter, said method comprising:

(a) contacting a library of compounds with a cell that expresses a recombinant anti-cell death gene and that comprises a reporter gene operably linked to said promoter, whereby expression of said reporter gene is altered if said library comprises a compound which modulates gene expression from said promoter;

(b) determining whether said reporter gene expression in said cell is altered as a result of contact with said compound library; and

(c) if said reporter gene expression is altered, identifying a compound from said library which modulates said reporter gene expression.

18. The method of claim 17, wherein step (c) comprises (i) dividing said library into two or more libraries with less complexity; and (ii) repeating steps (a) and (b) until a compound which modulates gene expression from a promoter is identified.

19. The method of claim 1, 3, 5, or 17, wherein said promoter is a heterologous promoter.

20. The method of claim 1, 3, 5, or 17, wherein said promoter is derived from a mammal and said library of polypeptides comprises polypeptides derived from a bacterium or a virus.

21. The method of claim 1, 3, 5, or 17, wherein said reporter gene is GFP.

22. The method of claim 1, 3, 5, or 17, wherein said anti-cell death gene is selected from the group consisting of bcl family members, IAP family members, and crmA.

23. The method of claim 1, 3, 5, or 17, wherein said cell is selected from the group consisting of CHO, CD-1, Cos, 293, HeLa, BHK, or L cells.

24. A method for determining whether a compound modulates NF-kB biological activity, said method comprising the steps of:

a) providing a cell expressing a BCMA polypeptide;

b) contacting said cell with a candidate compound; and

c) measuring the level of expression of said BCMA polypeptide in said cell, wherein a change in the level of expression of said BCMA polypeptide in said cell, relative to a cell not contacted with said candidate compound, identifies said candidate compound as a compound that modulates NF-kB biological activity.

25. A method for determining whether a compound modulates BCMA biological activity, said method comprising the steps of:

a) providing a BCMA polypeptide;

b) contacting said polypeptide with a candidate compound; and

c) measuring the level of biological activity of said BCMA polypeptide, wherein a change in the level of biological activity of said BCMA polypeptide, relative to a polypeptide not contacted with said candidate compound, identifies said candidate compound as a compound that modulates BCMA biological activity.

26. The method of claim 25, wherein said BCMA polypeptide is in a cell.

27. The method of claim 25, wherein said BCMA polypeptide is in a cell-free system.

28. The method of claim 25, wherein said BCMA biological activity is the modulation of NF-kB biological activity.

29. The method of claim 28, wherein said NF-kB biological activity is the modulation of cell death.

30. The method of claim 25, wherein said BCMA polypeptide comprises a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

31. A method for determining whether a compound modulates NF-kB activity, said method comprising the steps of:

a) providing a BCMA polypeptide;

b) contacting said polypeptide with a candidate compound; and

c) detecting the binding of said candidate compound to said polypeptide, wherein a candidate compound that binds to said polypeptide is a compound that modulates NF-kB biological activity.

32. The method of claim 31, wherein said BCMA polypeptide comprises a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) or amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

32. A substantially purified polypeptide comprising a polypeptide sequence having substantial identity to amino acids 98-164 of human BCMA (SEQ ID NO: 1) and not having amino acids 1-54 of human BCMA (SEQ ID NO: 1).

33. The polypeptide of claim 32, wherein said polypeptide modulates NF-kB activity.

34. The polypeptide of claim 32, wherein said polypeptide consists of amino acids 98-164 of human BCMA (SEQ ID NO: 1).

35. A substantially purified polypeptide comprising a polypeptide sequence having substantial identity to amino acids 97-163 of mouse BCMA (SEQ ID NO: 2) and not having amino acids 1-49 of mouse BCMA (SEQ ID NO: 2).

36. The polypeptide of claim 35, wherein said polypeptide

modulates NF-kB activity.

37. The polypeptide of claim 35, wherein said polypeptide consists of amino acids 97-163 of mouse BCMA (SEQ ID NO: 2).

38. An NF-kB modulator comprising a polypeptide of claim 32 or 35 covalently linked to a heterologous compound.

39. The modulator of claim 38, wherein said modulator modulates NF-kB activity.

40. A method for activating NF-kB activity in a cell, comprising contacting said cell with a recombinant BCMA polypeptide having NF-kB activating activity.

41. A method for activating NF-kB activity in a cell, comprising contacting said cell with a recombinant nucleic acid molecule encoding a BCMA polypeptide having NF-kB activating activity.

42. Use of a BCMA polypeptide for preparing a pharmaceutical composition for treating cancer, apoptosis, a viral infection, or an inflammatory response.

43. Use of a BCMA nucleic acid molecule for preparing a pharmaceutical composition for treating cancer, apoptosis, a viral infection, or an inflammatory response.

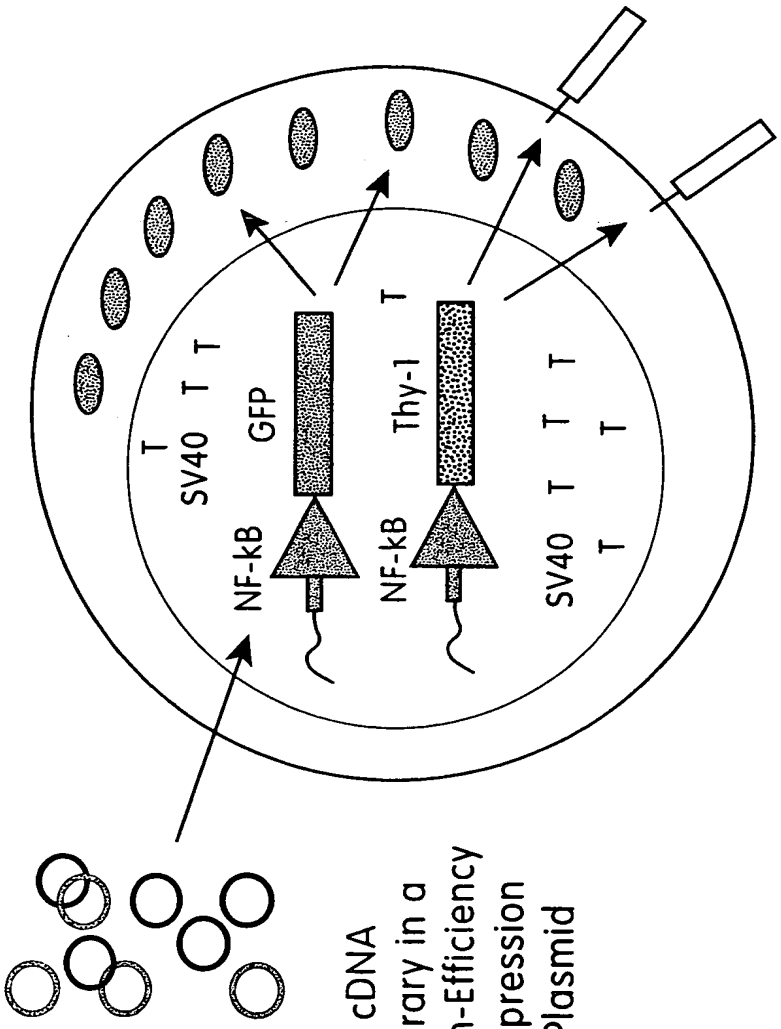

cDNA
Library in a
High-Efficiency
Expression
Plasmid
NF-kB
GFP
T
SV40 T T
NF-kB
Thy-1
T
SV40 T T
T
T

Fig. 1

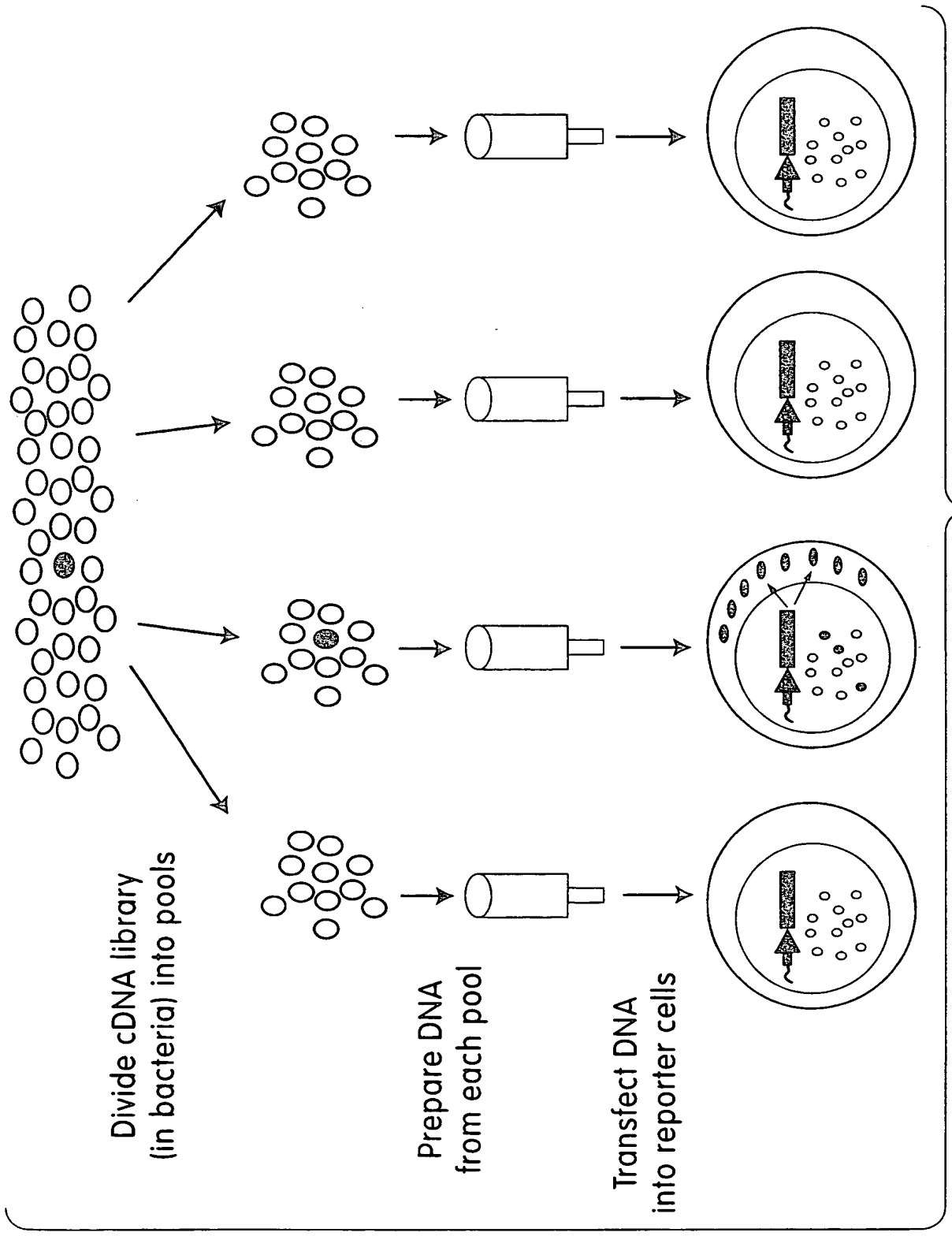

Divide cDNA library
(in bacteria) into pools
Prepare DNA
from each pool
Transfect DNA
into reporter cells

Fig. 2

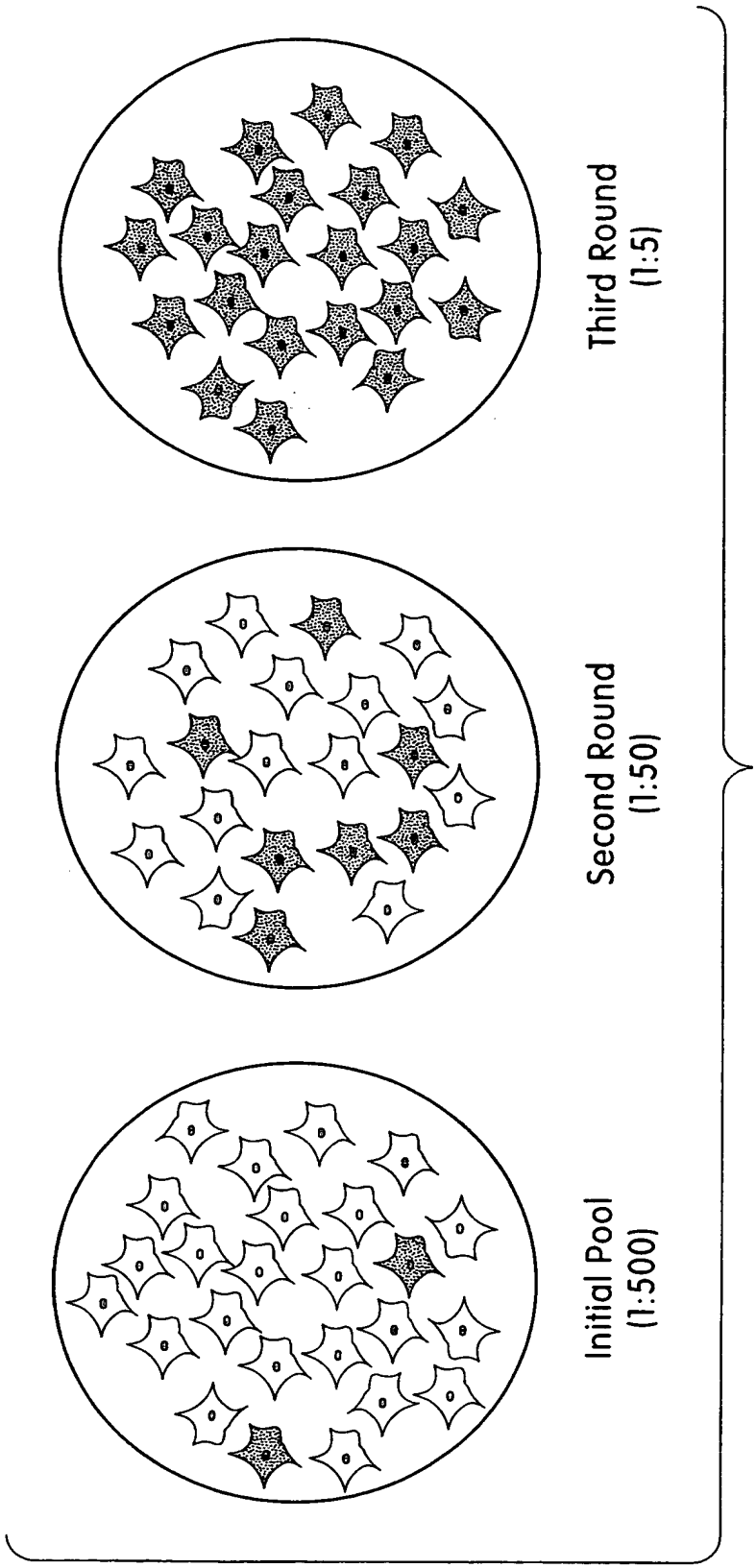


Fig. 3

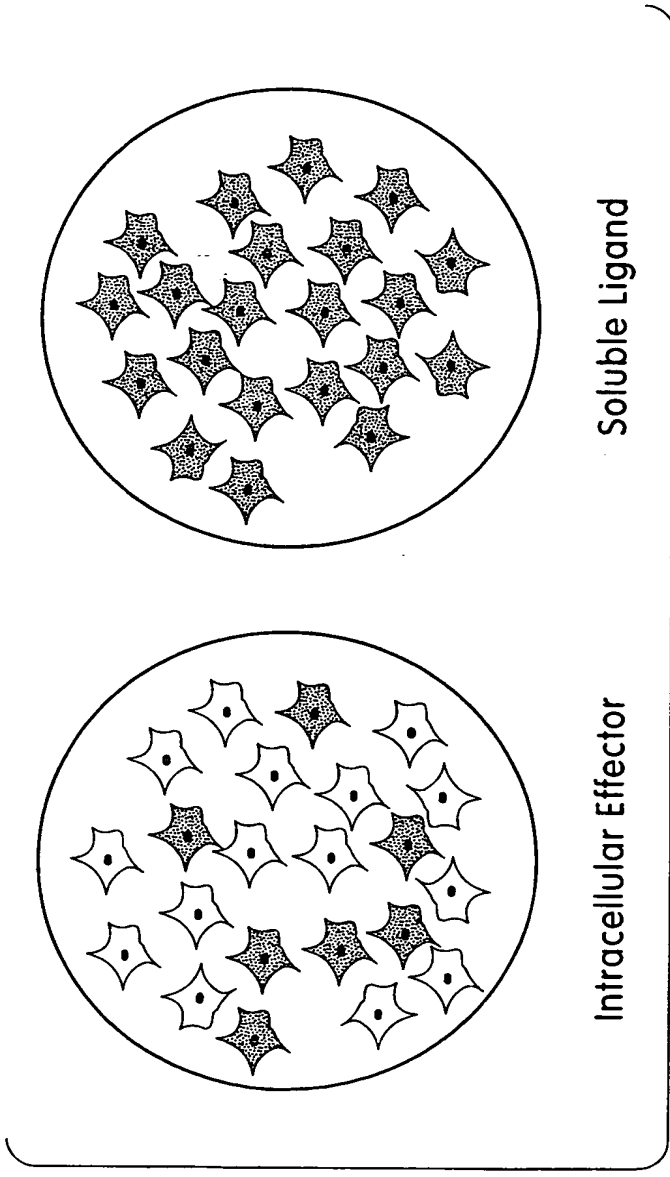


Fig. 4

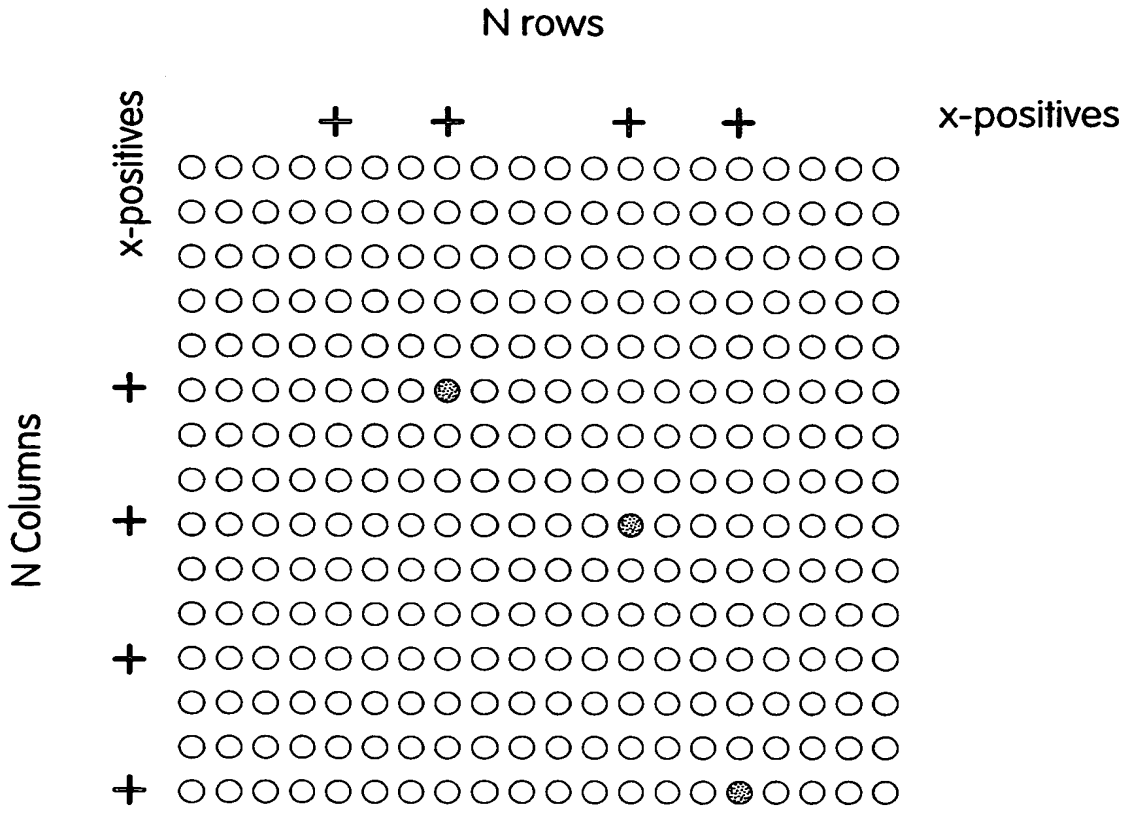


Fig. 5

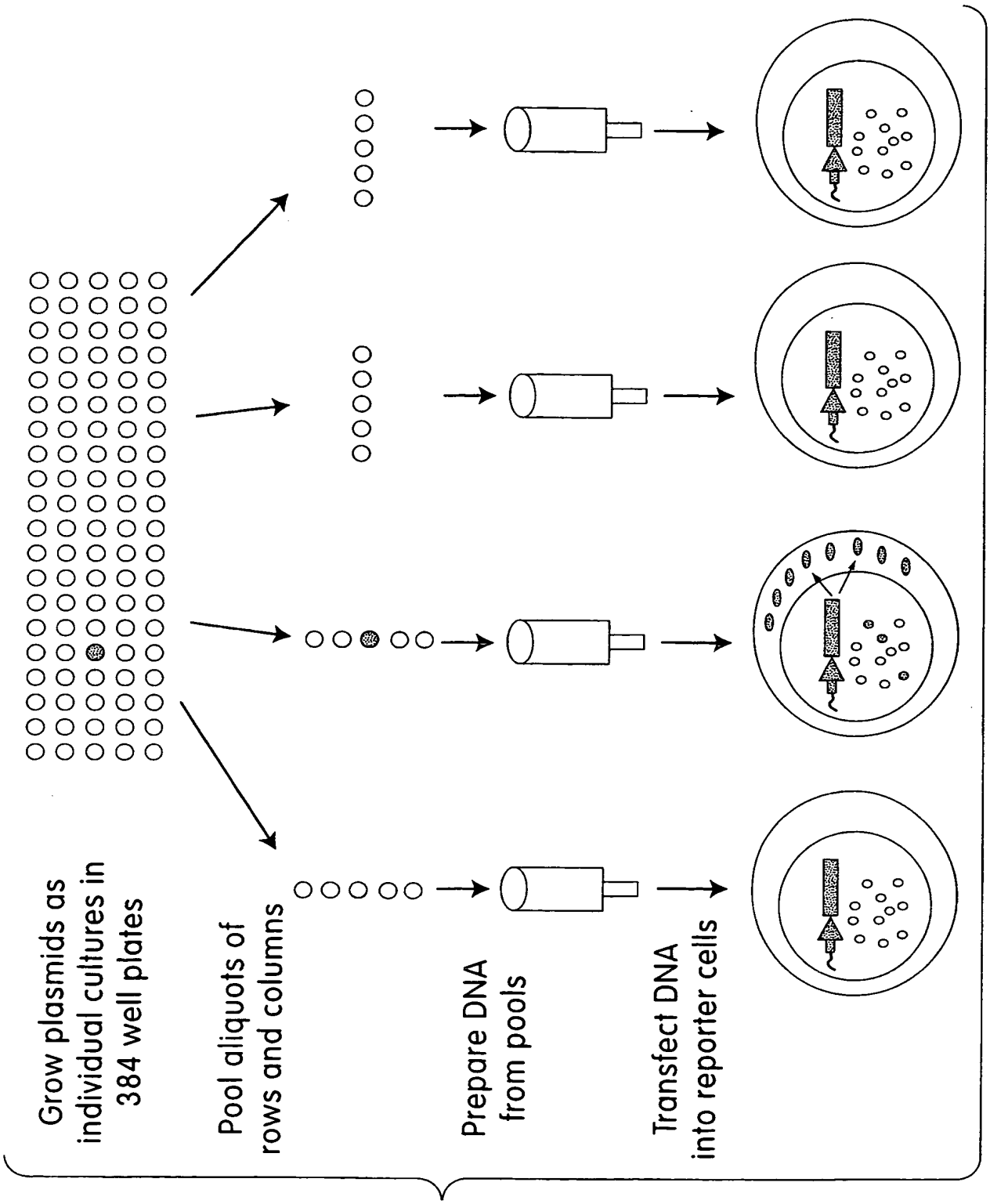

Grow plasmids as individual cultures in 384 well plates
Pool aliquots of rows and columns
Prepare DNA from pools
Transfect DNA into reporter cells

Fig. 6

Human BCNA

MLQMAGQCSQ NEYFDSLLHA CIPCQLRCSS NTPPLTCQRY CNASVTNSVK GTNA**ILWTCL**
GLSLIISLAV FVLMFLLRKI SSEPLKDEFK NTGSGLLGMA NIDLEKSRTG DEIILPRGLE
YTVEECTCED CIKSKPKVDS DHCFPLPAME EGATILVTTK TNDYCKSLPA ALSATEIEKS
ISAR (SEQ ID NO: 1)

Fig. 7A

Mouse BCNA

MAQQCFHSEY FDSLLHACKP CHLRCSNPPA TCQPYCDPSV TSSVKGTYT**V LWIFLGLTLV**
LSLALFTISF LLRKMNPEAL KDEPQSPGQL DGSAQLDKAD TELTRIRAGD DRIFPRSLEY
TVEECTCEDC VKSKPKGDSD HFFPLPAMEE GATILVTTKT GDYGKSSVPT ALQSVMGMEK
PTHTR (SEQ ID NO: 2)

Fig. 7B

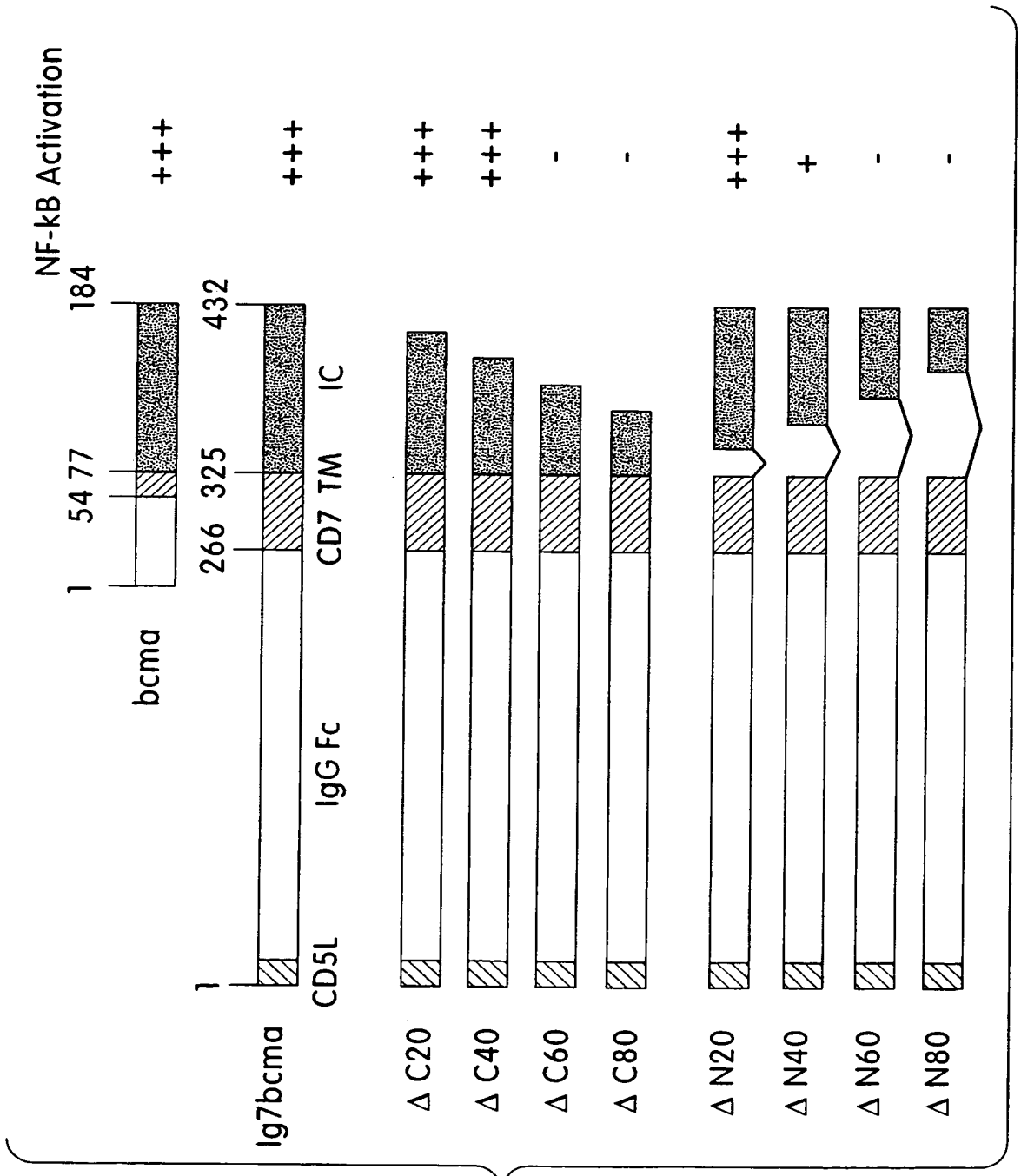

NF-kB Activation
bcma
1
54 77
184
Ig7bcma
1
CD5L
IgG Fc
266 325
432
CD7 TM
IC
Δ C20
Δ C40
Δ C60
Δ C80
Δ N20
Δ N40
Δ N60
Δ N80
+++
+++
+++
+++
-
-
+++
+
-
-

Fig. 8

SEQUENCE LISTING

<110> The General Hospital Corporation

<120> METHOD FOR CLONING SIGNAL TRANSDUCTION INTERMEDIATES

<130> 00786/371WO2

<150> 60/121,485
<151> 1999-02-24

<160> 2

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 184
<212> PRT
<213> Homo sapiens

<400> 1

```
Met Leu Gln Met Ala Gly Gln Cys Ser Gln Asn Glu Tyr Phe Asp Ser
 1               5                  10                  15
Leu Leu His Ala Cys Ile Pro Cys Gln Leu Arg Cys Ser Ser Asn Thr
            20                  25                  30
Pro Pro Leu Thr Cys Gln Arg Tyr Cys Asn Ala Ser Val Thr Asn Ser
        35                  40                  45
Val Lys Gly Thr Asn Ala Ile Leu Trp Thr Cys Leu Gly Leu Ser Leu
    50                  55                  60
Ile Ile Ser Leu Ala Val Phe Val Leu Met Phe Leu Leu Arg Lys Ile
65                  70                  75                  80
Ser Ser Glu Pro Leu Lys Asp Glu Phe Lys Asn Thr Gly Ser Gly Leu
                85                  90                  95
Leu Gly Met Ala Asn Ile Asp Leu Glu Lys Ser Arg Thr Gly Asp Glu
            100                 105                 110
Ile Ile Leu Pro Arg Gly Leu Glu Tyr Thr Val Glu Glu Cys Thr Cys
        115                 120                 125
Glu Asp Cys Ile Lys Ser Lys Pro Lys Val Asp Ser Asp His Cys Phe
    130                 135                 140
Pro Leu Pro Ala Met Glu Glu Gly Ala Thr Ile Leu Val Thr Thr Lys
145                 150                 155                 160
Thr Asn Asp Tyr Cys Lys Ser Leu Pro Ala Ala Leu Ser Ala Thr Glu
                165                 170                 175
Ile Glu Lys Ser Ile Ser Ala Arg
            180
```

<210> 2
<211> 185
<212> PRT
<213> Mus musculus

<400> 2

```
Met Ala Gln Gln Cys Phe His Ser Glu Tyr Phe Asp Ser Leu Leu His
```

2

```
 1               5                  10                  15
Ala Cys Lys Pro Cys His Leu Arg Cys Ser Asn Pro Pro Ala Thr Cys
            20                  25                  30
Gln Pro Tyr Cys Asp Pro Ser Val Thr Ser Ser Val Lys Gly Thr Tyr
        35                  40                  45
Thr Val Leu Trp Ile Phe Leu Gly Leu Thr Leu Val Leu Ser Leu Ala
    50                  55                  60
Leu Phe Thr Ile Ser Phe Leu Leu Arg Lys Met Asn Pro Glu Ala Leu
65                  70                  75                  80
Lys Asp Glu Pro Gln Ser Pro Gly Gln Leu Asp Gly Ser Ala Gln Leu
                85                  90                  95
Asp Lys Ala Asp Thr Glu Leu Thr Arg Ile Arg Ala Gly Asp Asp Arg
            100                 105                 110
Ile Phe Pro Arg Ser Leu Glu Tyr Thr Val Glu Glu Cys Thr Cys Glu
        115                 120                 125
Asp Cys Val Lys Ser Lys Pro Lys Gly Asp Ser Asp His Phe Phe Pro
    130                 135                 140
Leu Pro Ala Met Glu Glu Gly Ala Thr Ile Leu Val Thr Thr Lys Thr
145                 150                 155                 160
Gly Asp Tyr Gly Lys Ser Ser Val Pro Thr Ala Leu Gln Ser Val Met
                165                 170                 175
Gly Met Glu Lys Pro Thr His Thr Arg
            180                 185
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WEST 1.2. MEDLINE, SCISEARCH, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,874,304 A (ZOLOTUKHIN et al) 23 February 1999(23.02.99), see the entire document.	1-4, 7-16, 19-23
Y	CLEM, R.J et al. Prevention of Apoptosis by a Baculovirus Gene During Infection of Insect Cells. Science. 29 November 1991, Vol. 254, pages 1388-1390, especially page 1388.	1-4, 7-16, 19-23
Y	BIRNBAUM et al. An Apoptosis-inhibiting Gene from a Nuclear Polyhedrosis Virus Encoding a Polypeptide with a Cys/His Sequence Motifs. Journal of Virology. April 1994, Vol. 68, No. 4, pages 2521-2528, especially pages 2523-2524.	1-4, 7-16, 19-23

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 10 MAY 2000	Date of mailing of the international search report 19 JUL 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer P. PONNALURI Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998) *

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PASARELLI et al. A Baculovirus Gene Involved in Late Gene Expression Predicts a Large Polypeptide with a Conserved Motif of RNA Polymerases. Jounal of Virology. July 1994, vol. 68, No. 7, pages 4673-4678, see entire document.	1-4, 7-16, 19-23

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4, 7-16, 19-23

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/04925

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C12Q 1/00, 1/70, 1/66; C12N 7/00, 5/00, 5/02; G01N 33/554; A61K 38/00; C07H 21/02, 21/04, 21/00

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/4, 5, 8, 235.1, 325; 436/518; 530/300; 536/23.1, 23.4, 23.7, 25.32; 935/90, 93, 95, 106

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-4, 7-16, 19-23, drawn to a method for identifying a polypeptide which increases gene expression from a promoter.
Group II, claim(s) 5-16, 19-23, drawn to a method of identifying a polypeptide which modulates activation of a transcription factor activation domain.
Group III, claim(s) 17-23, drawn to a method of identifying a compound which modulates gene expression from a promoter.
Group IV, claim(s) 24, drawn to a method of determining whether a compound modulates NF-kB biological activity.
Group V, claim(s) 25-30, drawn to a method of determining a compound which modulates BCMA biological activity.
Group VI, claim(s) 31-32, drawn to a method of determining a compound NF-kB activity.
Group VII, claim(s) 33-37, drawn to a polypeptide which modulates NF-kB activity.
Group VIII, claim(s) 38-39, drawn to an NF-kB modulator.
Group IX, claim(s) 40-41, drawn to a method of activating NF-kB activity in a cell.
Group X, claim(s) 42-43, drawn to use of a BCMA polypeptide or nucleic acid.
The inventions listed as Groups I-X do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical feature of group I is polypeptides which increase gene expression is not present in other groups, and also the polypeptides which increase gene expression from a promoter and method of identifying is known. Thus, the groups lack unity.